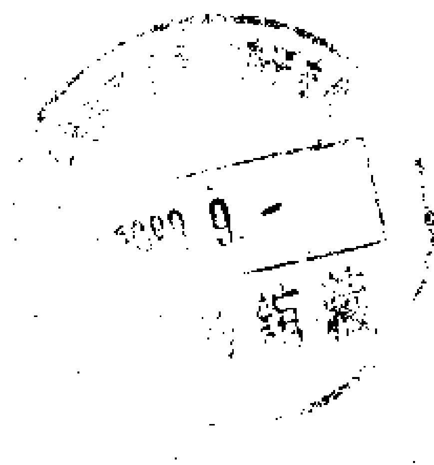
流形上的微积分

欧阳光中



上海科学技术出版社

内 容 简 介

本书以现代数学的观点处理数学分析中的许多经典内容，并加以拓广到流形上。微分流形是三维欧氏空间中的光滑曲线(一维微分流形)和光滑曲面(二维微分流形)在高维空间甚至一般拓扑空间中自然而直接的拓广；流形上的积分是曲线积分和曲面积分的拓广。在本书中，把经典数学分析中的几个著名公式，如格林公式、高斯公式、斯托克司公式等在高维的流形上，利用外微分，统一为一个形式、主要内容包括：欧氏空间上映射的微分学和积分学，微分流形，微分形式和外微分，流形上的积分，斯托克司公式等，本书可作为理工科大学的教材。供二年级下学期或三年级上学期使用；也可作为学过普通微积分和线性代数的大学生的参考书和课外读物。

流形上的微积分
欧阳光中 编
责任编辑 赵序明
上海科学技术出版社出版
(上海瑞金二路450号)
新华书店上海发行所发行 祝桥新华印刷厂印刷
开本850×1156 1/32 印张5.75 字数125,000
1988年3月第1版 1988年3月第1次印刷
印数1—4,500
ISBN 7-5323-0018-8/O·1
定价：2.10元

前　　言

这本书是写给已经读过通常的数学分析和线性代数的大学生们看的，目的是向读者介绍现代分析中的一个基础内容，并使读者能够运用现代数学的观点，从更高的层次上回过头来看待曾经学习过的许多非常重要的经典理论，这对从事数学、应用数学、计算数学、物理、力学等方面的教学和研究都是十分必要的。

正是由于上述原因，当前不少理工科大学和师范院校为学生开设了流形上的微积分这门课．本书可作为这门课程的教材；另一方面，据了解，还有许多担任数学基础课的教师也希望借鉴这方面的知识来改革基础课的教学内容，以进一步提高教学质量．本书也适合于这些教师和学生。

读者在开始阅读本书时，如果想预先了解本书的基本内容是什么，建议读者先阅读本书每一章节的开头部分，在那里指明了每一章节的中心要点。

本书的编写，对几乎每一个重要的概念和理论都尽可能讲明它的来龙去脉，指出了它是怎样思考和发现的．这样，对读者阅读本书可能有较多的帮助。

限于作者水平，本书中不当之处在所难免，恳切希望读者提出批评指正。

欧阳光中

目　录

第1章　欧几里得空间

§1.1　欧几里得空间上的范数

本节的中心问题是：从通常的数学分析中邻域的概念出发进行讨论．以平面为例，设点 $M_0=(x_0, y_0)$ 是平面上的一点，以 M_0 为中心既可以作一个圆形邻域 $O(M_0, \delta)$，也可以作一个方形邻域 $O'(M_0, \delta)$，它们的定义分别是

$$O(M_0, \delta)=\{(x, y)\in R^2 \mid (x-x_0)^2+(y-y_0)^2<\delta^2\};$$

$$O'(M_0, \delta)=\{(x, y)\in R^2 \mid |x-x_0|<\delta, |y-y_0|<\delta\}.$$

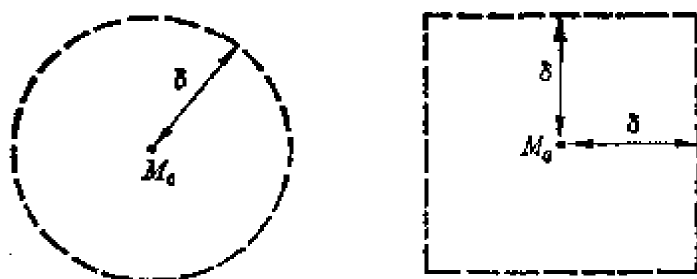


图　1-1

（图1-1画出了这两个邻域）．平面上的一个点列 $M_n=(x_n, y_n)$ $(n=1, 2, \cdots)$ 收敛于点 M_0，可以用两种方式来陈述：

(i) 圆形邻域方式：对任意给定的 $\varepsilon>0$，存在正整数 N，当 $n>N$ 时，有 $\sqrt{(x_n-x_0)^2+(y_n-y_0)^2}<\varepsilon$．或者说，当 n 充分大时，点列 $M_n\in O(M_0, \varepsilon)$．

(ii) 方形邻域方式：对任意给定的 $\varepsilon>0$，存在正整数 N，当 $n>N$ 时，有 $|x_n-x_0|<\varepsilon$，$|y_n-y_0|<\varepsilon$．或者说，当 n 充分大时，点列 $M_n\in O'(M_0, \varepsilon)$．

这两种陈述方式是相互一致的，也就是说，在一种方式下是收敛的，在另一种方式下也一定是收敛的．本节要把这一事实作非

常一般的论述。

欧几里得范数

设 R^n 是 n 维（实）欧几里得空间，其中的点（或者称它是向量）x 可以用 $(x_1, x_2, \cdots, x_n)$ 表示，$x_i(i=1, 2, \cdots, n)$ 是实数，称它是 x 的 i-坐标。当我们说 R^n 是一个（实）欧几里得空间的时候，意味着 R^n 中任何两点 $x=(x_1, \cdots, x_n)$、$y=(y_1, \cdots, y_n)$ 可以作出它们的和 $x+y=(x_1+y_1, \cdots, x_n+y_n)\in R^n$，还可以用任一实数 α 乘向量 x 得到 $\alpha x=(\alpha x_1, \cdots, \alpha x_n)\in R^n$，并且这两种运算满足若干基本的法则，使得点集 $R^n=\{x=(x_1, \cdots, x_n)\,|\,x_i$ 是实数，$i=1, 2, \cdots, n\}$ 带有这两种运算成为一个实数域上的线性空间。不仅如此，在 R^n 中还带有内积

$$\langle x, y\rangle = x_1y_1+\cdots+x_ny_n.$$

由内积的定义不难验证它具有以下基本特性

(i) 对称性： $\langle x, y\rangle=\langle y, x\rangle$, $\forall x, y\in R^n$.

(ii) 线性： $\langle x+y, z\rangle=\langle x, z\rangle+\langle y, z\rangle$,

$\langle \alpha x, y\rangle=\alpha\langle x, y\rangle$, $\forall x, y, z\in R^n$; $\forall \alpha\in R$（实数）.

(iii) 正定性： $\langle x, x\rangle\geqslant 0$, $\forall x\in R^n$.

$\langle x, x\rangle=0$ 当且仅当 $x=0$（这里 0 是零向量）.

由(i)和(ii)立即得出

$$\langle x, y+z\rangle=\langle x, y\rangle+\langle x, z\rangle,$$

$$\langle x, \alpha y\rangle=\alpha\langle x, y\rangle, \quad \forall x, y, z\in R^n; \forall\alpha\in R.$$

从内积出发，可以产生向量 x 的模长 $|x|$：

$$|x|=\sqrt{\langle x, x\rangle}=\sqrt{x_1^2+\cdots+x_n^2}.$$

称它是 x 的欧几里得范数。欧几里得范数具有以下三个基本特性：

(i) $|x|\geqslant 0$, $\forall x\in R^n$.

$|x|=0$ 当且仅当 $x=0$.

(ii) $|\alpha x|=|\alpha|\cdot|x|$. $\forall x\in R^n$; $\forall\alpha\in R$.

(iii) $|x+y|\leqslant|x|+|y|$, $\forall x, y\in R^n$.

最后的不等式又称为三角形不等式，它的几何解释是：如果将向量 x 和 y 看作三角形的两个边，那么 $x+y$ 就是第三边。不等式表明三角形两边长度之和不小于第三边的长度。

(i) 和 (ii) 的证明是明显的。为了证明 (iii)，我们运用一个著名的许瓦兹(Schwarz)不等式。

引理(许瓦兹不等式)　对任何 $x、y\in R^n$，有

$$|\langle x, y\rangle|\leqslant|x|\cdot|y|.$$

证明　当 $y=0$ (零向量)时，不等式显然成立。

现设 $y\neq 0$，对任一实数 α，由内积的基本性质，有

$$0\leqslant\langle x-\alpha y, x-\alpha y\rangle=\langle x, x\rangle+\alpha^2\langle y, y\rangle-2\alpha\langle x, y\rangle,$$

令 $\alpha=\dfrac{\langle x, y\rangle}{\langle y, y\rangle}$，代入上式右端，得

$$\begin{aligned}&\langle x, x\rangle+\frac{\langle x, y\rangle^2}{\langle y, y\rangle}-2\frac{\langle x, y\rangle^2}{\langle y, y\rangle}\\&\quad=\frac{\langle x, x\rangle\langle y, y\rangle-\langle x, y\rangle^2}{\langle y, y\rangle}\geqslant 0,\end{aligned}$$

而分母 $\langle y, y\rangle>0$，这样就得出许瓦兹不等式。证毕。

应注意的是：在引理的证明中只利用了内积的三条基本特性，并没有利用点 x 和 y 在 R^n 中的坐标表示，也没有涉及空间的维数，这就启发我们考虑这样一个事实：许瓦兹不等式不限于有限维的欧几里得空间。确实如此，在比欧几里得空间更一般的具有内积的线性空间内，许瓦兹不等式也成立。

现在回过头来证明三角形不等式

$$|x+y|\leqslant|x|+|y|,\quad\forall x, y\in R^n.$$

证明

$$\begin{aligned}|x+y|^2&=\langle x+y, x+y\rangle=\langle x, x\rangle+2\langle x, y\rangle+\langle y, y\rangle\\&=|x|^2+2\langle x, y\rangle+|y|^2,\end{aligned}$$

利用许瓦兹不等式

$$\begin{aligned}|x+y|^2&\leqslant|x|^2+2|x|\cdot|y|+|y|^2\\&=(|x|+|y|)^2.\end{aligned}$$

证毕

在这个证明中如同在许瓦兹不等式的证明中一样，没有涉及

到点的坐标和空间的维数，因此，它也具有某种更普遍的意义。

下面，我们将从欧几里得范数的三个最基本的特性出发，加以抽象化，以此引进一般的范数的概念。

范数

我们暂时离开欧几里得空间，考察一个更广的空间，即线性空间，然后带着从更一般的空间中所获得的观点，再回到欧几里得空间中深入的讨论有关问题。

设X是一个线性空间。当我们说X是一个线性空间的时候，意味着X内有“加法”运算$x+y$(其中x、$y\in X$)和“数乘”运算αx(其中$x\in X$, $\alpha\in R$)，它们服从必要的法则。在“数乘”中，α一般是复数，在本书中，我们限定α是实数，考虑(实)线性空间。它可以是有限维的，也可以是无限维的。

在X上定义一个映射$\|\ \|$($\|\ \|$是映射的记号)

$$\|\ \|: X\to R$$
$$x\mapsto \|x\|,$$

即对每一点$x\in X$，其象$\|x\|$是一个实数。如果这一映射满足下列三个条件：

(i) $\|x\|\geqslant 0$, $\quad\forall x\in X$.

$\|x\|=0$当且仅当$x=0$(右端的0是X中的零元)。

(ii) $\|\alpha x\|=|\alpha|\|x\|$, $\quad\forall x\in X, \forall\alpha\in R$.

(iii) $\|x+y\|\leqslant\|x\|+\|y\|$, $\quad\forall x$、$y\in X$.(即三角形不等式)。

称这个映射$\|\ \|$是X上的一个范数。又称$\|x\|$是x的范数。当线性空间X上具备了范数以后，就称X是一个赋范线性空间，通常用$(X,\|\ \|)$表示它。

例如欧几里得范数就是R^n上的一个范数。

在现代分析中，范数有什么作用呢？我们先考察一下欧几里得范数所起的作用。在空间R^n中，由欧几里得范数可以定义距离，任何两点$x=(x_1, \cdots, x_n)\in R^n$, $y=(y_1, \cdots, y_n)\in R^n$的距离就是$x-y$的范数$\|x-y\|$，即

$$|x-y|=\sqrt{(x_1-y_1)^2+\cdots+(x_n-y_n)^2}.$$

有了距离就能够正确的表达何谓收敛，还可以引进开球，例如

$$O(a,\ \delta)=\{x\in R^n \mid |x-a|<\delta\}$$

就是一个以点 $a\in R^n$ 为中心，以 δ 为半径的开球．从而引进内点，外点，边界点，开集，闭集，聚点，紧集，连通等一系列重要的概念，整个分析"大厦"就是建筑在这些概念基础之上的。

范数在线性空间中所起的作用犹如欧几里得范数在 R^n 中的作用一样，设 $(x,\ \|\ \ \|)$ 是一个赋范线性空间，和欧几里得空间相仿的——几乎是照抄——可以获得下述基本概念。

距离　X 内任何两点 x 和 y，定义它们的距离 $\rho(x,\ y)$ 是

$$\rho(x,\ y)=\|x-y\|.$$

由范数满足的条件容易知道距离具有以下性质：

(i) $\rho(x,\ y)=\|x-y\|\geqslant 0,\quad \forall x、y\in X,$

$$\rho(x,\ y)=0 \text{ 当且仅当 } x=y.$$

(ii) $\rho(x,\ y)\leqslant\rho(x,\ z)+\rho(z,\ y),\quad \forall x、y、z\in X.$

(iii) $\rho(x, y)=\rho(y,\ x),\quad \forall x, y\in X.$

开球　对 X 中的任何一点 a 和任何实数 $\delta>0$，称

$$O(a,\ \delta)=\{x\in X \mid \|x-a\|<\delta\}$$

是以点 a 为中心，以 δ 为半径的开球，又称它是点 a 的球邻域。

收敛　设 $x_n\in X\,(n=1,\ 2,\ \cdots)$，$x\in X$，如果对任意给定的 $\varepsilon>0$，存在正整数 N，当 $n>N$ 时，有

$$\|x_n-x\|<\varepsilon \text{ 或 } x_n\in O(x,\ \varepsilon),$$

则称点列 $\{x_n\}$ 收敛于 x，记为 $\lim\limits_{n\to\infty}x_n=x$ 或 $x_n\to x\,(n\to\infty)$。

开集　设 $S\subset X$，$x\in S$，如果存在 x 的一个球邻域 $O(x,\ \delta)\subset S$，则称点 x 是 S 的一个内点。如果 S 的点都是 S 的内点，则称 S 是开集。

闭集　设 $S\subset X$，如果 S 的补集 $S^C=X-S$ 是开集，则称 S 是闭集。

此外如聚点，闭包，外点，边界点和边界，紧集，连通等概念都

可几乎一字不差的照抄欧几里得空间中相应的概念，这里不一一列举了。在本章中我们涉及的是欧几里得空间，因此也没有必要深入讨论一般的赋范线性空间。

例 1 设 $C[a, b]$ 是闭区间 $[a, b]$ 上的所有实值连续函数组成的线性空间，其中的“加”和“数乘”的定义为

$$(f+g)(x)=f(x)+g(x),$$
$$(\alpha f)(x)=\alpha f(x), \quad x\in[a, b].$$

在 $C[a,b]$ 上装备如下的范数：

$$\|f\|=\sup\{|f(x)|\,\big|\,a\leqslant x\leqslant b\}, \quad f\in C[a, b].$$

容易验证 $\|\ \|$ 满足范数的三个条件，它确实是 $C[a, b]$ 上的一个范数。现在研究在这一范数的意义下收敛将意味着什么。

设 $f_n\in C[a, b]$，$f\in C[a, b]$，如果对任意给定的 $\varepsilon>0$，存在正整数 N，当 $n>N$ 时，有 $\|f_n-f\|<\varepsilon$，就称在范数 $\|\ \|$ 下 f_n 收敛于 f，记为 $f_n\to f$，或者记为 $\|f_n-f\|\to 0(n\to\infty)$。这也就是

$$\|f_n-f\|=\sup\{|f_n(x)-f(x)|\,\big|\,a\leqslant x\leqslant b\}\to 0,$$

它正是数学分析中一致收敛的概念。换句话说，在 $[a, b]$ 上 f_n 一致收敛于 f 就是在上述范数的意义下 f_n 收敛于 f。

例 2 $C[a, b]$ 如同例 1 所述。在 $C[a, b]$ 上装备另外一个范数。

$$\|f\|_1=\int_a^b|f(x)|dx, \quad f\in C[a, b].$$

不难验证 $\|\ \|_1$ 是 $C[a, b]$ 上的范数。

在这一范数的意义下，$C[a, b]$ 上序列 f_n 收敛于 $f\in C[a, b]$，就是

$$\|f_n-f\|_1=\int_a^b|f_n(x)-f(x)|dx\to 0 \quad (n\to\infty).$$

对读者来说这也并不陌生，它正是数学分析中平均收敛的概念。换句话说 $C[a, b]$ 上的序列 f_n 平均收敛于 $f\in C[a, b]$，就是在上述范数的意义下 f_n 收敛于 f。

例 1 和例 2 是在同一个线性空间 $C[a, b]$ 上装备了两个范数

$\| \|$和$\| \|_1$，形成两个赋范线性空间$(C[a, b], \| \|)$和$(C[a, b], \| \|_1)$，在这两个空间内都有各自的收敛概念，这两种收敛之间有什么关系呢？下面有

(1) 设$f_n \in C[a, b]$，$f \in C[a, b]$。如果在范数$\| \|$的意义下$f_n \to f$，那么在范数$\| \|_1$的意义下也有$f_n \to f$。

(2) 存在序列$f_n \in C[a, b]$，$f \in C[a, b]$，在范数$\| \|_1$的意义下$f_n \to f$，但在范数$\| \|$的意义下f_n不收敛于f。

第一个结论由

$$\begin{aligned}\|f_n - f\|_1 &= \int_a^b |f_n(x) - f(x)|\,dx \\ &\leqslant \sup\{|f_n(x) - f(x)| \,\big|\, a \leqslant x \leqslant b\} \cdot \int_a^b dx \\ &= \|f_n - f\| \cdot (b-a)\end{aligned}$$

立即得出。

第二个结论可以考虑相当直观的例子。设(图 1-2)

$$f_n(x) = \begin{cases} 2n^{3/2}x, & 0 \leqslant x < \dfrac{1}{2n}; \\ -2n^{3/2}\left(x - \dfrac{1}{n}\right), & \dfrac{1}{2n} \leqslant x < \dfrac{1}{n}; \\ 0, & \dfrac{1}{n} \leqslant x \leqslant 1. \end{cases}$$

$$f(x) = 0, \quad 0 \leqslant x \leqslant 1.$$

$f_n \in C[0, 1]$，$f \in C[0, 1]$，并且

$$\begin{aligned}\|f_n - f\|_1 &= \int_0^1 |f_n(x) - f(x)|\,dx \\ &= \int_0^1 f_n(x)\,dx \\ &= \frac{1}{2\sqrt{n}} \to 0 \quad (n \to \infty)\end{aligned}$$

即在范数$\| \|_1$之下，f_n收敛于f。但

$$\begin{aligned}\|f_n - f\| &= \sup\{|f_n(x) - f(x)| \,|\, a \leqslant x \leqslant b\} \\ &= \sqrt{n} \to \infty \quad (n \to \infty).\end{aligned}$$

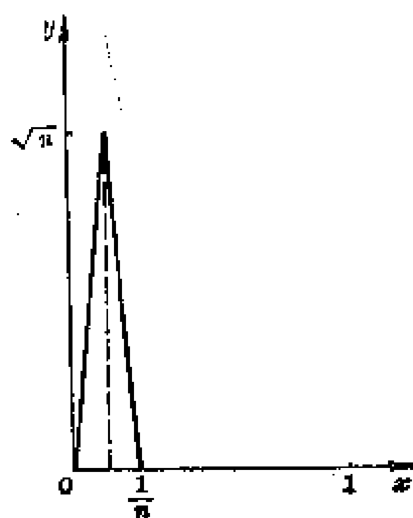


图 1-2

这说明两个范数$\| \ \|$和$\| \ \|_1$是不同的，因此我们说，$(C[a, b], \| \ \|)$和$(C[a, b], \| \ \|_1)$是两个不同的赋范线性空间。一般说来，在同一个线性空间X上装备了两个不同的范数$\| \ \|$和$\| \ \|'$，那么作为赋范线性空间$(X, \| \ \|)$和$(X, \| \ \|')$，它们是相异的。

例 3 设M是所有$m \times n$矩阵组成的集合，带有通常的矩阵的"加"和"数乘"，即设$A = (a_{ij})$，$B = (b_{ij})$ $(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$是两个$m \times n$矩阵，$\alpha$是一个实数，那么

$$A + B = (a_{ij} + b_{ij}), \quad \alpha A = (\alpha a_{ij}).$$

再定义M中的范数是

$$\|A\| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2},$$

这时M就成为一个赋范线性空间。

再设$A_k \in M$，$A \in M$，$A_k = (a_{ij}^{(k)})$，$A = (a_{ij})$，容易看出$A_k \to A$ (即$\|A_k - A\| \to 0$)等价于

$$a_{ij}^{(k)} \to a_{ij}, \ (k \to \infty; i = 1, 2, \cdots, m; j = 1, 2, \cdots, n).$$

例 4 在 2 维欧几里得空间R^2内，对每一点$x = (x_1, x_2) \in R^2$，定义三个范数（都不难验证它们确实满足范数的三条要求），

$$|x| = \sqrt{x_1^2 + x_2^2} \quad \text{(欧几里得范数)}$$

$$\|x\|_{(1)} = |x_1| + |x_2|,$$

$$\|x\|_{(2)} = \max(|x_1|, |x_2|).$$

例如$x = (2, -3)$，那么$|x| = \sqrt{13}$，$\|x\|_{(1)} = 5$，$\|x\|_{(2)} = 3$。

在欧几里得范数$| \ |$的意义下单位开球是(图 1-3(a))

$$O(0,1) = \{x \in R^2 \mid |x| < 1\}$$
$$= \{x \in R^2 \mid x_1^2 + x_2^2 < 1\}.$$

在范数$\| \ \|_{(1)}$的意义下单位开球是(图 1-3(b))

$$O(0, 1) = \{x \in R^2 \mid \|x\|_{(1)} < 1\}$$
$$= \{x \in R^2 \mid |x_1| + |x_2| < 1\}.$$

在范数$\| \ \|_{(2)}$的意义下单位开球是(图 1-3(c))

$$\mathcal{O}(0, 1) = \{x \in R^2 \mid \|x\|_{(2)} < 1\}$$
$$= \{x \in R^2 \mid |x_1| < 1, |x_2| < 1\}.$$

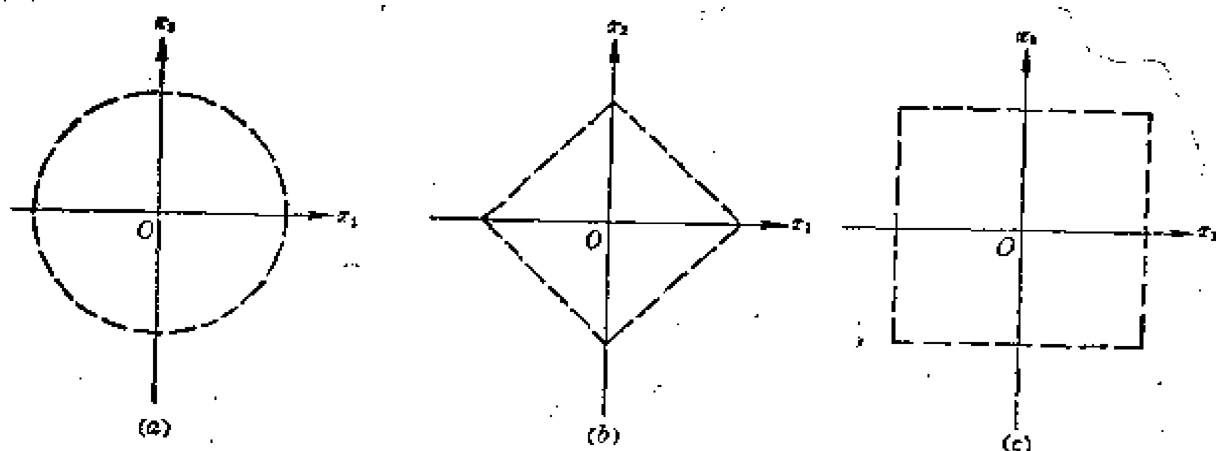


图 1-3

从直观上不难看出，在这三个范数意义下的收敛是相同的，也就是说，在其中任何一个范数下是收敛的，那么在其他两个范数下一定也是收敛的。现在的问题是：在 R^2 内可以引进各种形式上不同的范数，在这些范数下收敛是否相同呢？是否也会出现例 1 和例 2 的情形呢？

欧几里得空间上范数的等价性

设在同一个线性空间 X 上装备了两个范数 $\|\ \|_1$ 和 $\|\ \|_2$，如果存在两个常数 $c>0$，$d>0$，使

$$\|x\|_1 \leqslant c\|x\|_2, \quad \|x\|_2 \leqslant d\|x\|_1, \quad \forall x \in X.$$

就称范数 $\|\ \|_1$ 和 $\|\ \|_2$ 是等价的范数。

在赋范线性空间 $(X, \|\ \|_1)$ 和 $(X, \|\ \|_2)$ 中，如果 $\|\ \|_1$ 和 $\|\ \|_2$ 是等价的范数，那么在这两个空间中收敛是相同的，即在范数 $\|\ \|_1$ 的意义下 $x_n \to x$ 当且仅当在范数 $\|\ \|_2$ 的意义下 $x_n \to x$。例 1 和例 2 告诉我们线性空间 $C[a, b]$ 上的两个范数

$$\|f\| = \sup\{|f(x)| \mid x \in [a, b]\}$$

与

$$\|f\|_1 = \int_a^b |f(x)|\,dx$$

是不等价的。

一般说来，对任意一个线性空间，如果在它上面装备了两个范数，我们不能断言它们一定是等价的。然而对欧几里得空间而言，任何两个范数是等价的，这就是下面的定理。

定理（欧几里得空间上范数的等价性） 设$|\ |$是n维欧几里得空间R^n上的欧几里得范数，$\|\ \|$是R^n上任意一个范数，那么$|\ |$和$\|\ \|$等价。

证明 先证明存在常数$c>0$，使

$$\|x\|\leqslant c|x|,\quad x\in R^n.$$

设$e_1,e_2,\cdots,e_n$是R^n中的一组基，在这组基下每一个$x\in R^n$可以表示为

$$x=x_1e_1+\cdots+x_ne_n,\quad x_i\in R\quad(i=1,2,\cdots,n).$$

由范数所满足的条件知道

$$\|x\|\leqslant|x_1|\|e_1\|+\cdots+|x_n|\|e_n\|,$$

再利用$|x_i|\leqslant|x|\ (i=1,2,\cdots,n)$，得

$$\|x\|\leqslant(\|e_1\|+\cdots+\|e_n\|)|x|,$$

记$c=\|e_1\|+\cdots+\|e_n\|$，注意到$\|e_i\|>0(i=1,2,\cdots,n)$。这样便证明了

$$\|x\|\leqslant c|x|,\quad x\in R^n.$$

再证明存在常数$d>0$，使

$$|x|\leqslant d\|x\|,\quad x\in R^n,$$

考虑n元函数f

$$f:R^n\to R,$$

它的定义是

$$f(x)=\|x\|,\quad x\in R^n,$$

利用范数的三角形不等式，有

$$|f(x')-f(x)|=|\|x'\|-\|x\||\leqslant\|x'-x\|.$$

再由刚才已经证明的结论，上式右端不超过$c|x'-x|$($c>0$是一个常数)。这表明f是一个n元的连续函数。

作单位球面

$$S=\{x\in R^n \mid \|x\|=1\},$$

它是 R^n 中的一个有界闭集（也就是说它是 R^n 中的一个紧集），f 在 S 连续，所以 f 在 S 上必有最小值，设它是 α，即

$$\alpha=\min\{\|x\| \mid x\in S\}.$$

又因为对每一个 $x\in S, f(x)=\|x\|>0$，所以 $\alpha>0$。

设 x 是 R^n 中的任一向量，$x\neq 0$，那么 $\frac{x}{|x|}\in S$，所以

$$\alpha\leqslant f\left(\frac{x}{|x|}\right)=\left\|\frac{x}{|x|}\right\|=\frac{1}{|x|}\|x\|,$$

这就是

$$|x|\leqslant\frac{1}{\alpha}\|x\|,$$

当 $x=0$ 时，上式显然成立。再令 $d=\frac{1}{\alpha}$，定理证毕。

由定理立即获得下面的推论。

推论 1 欧几里得空间上任何两个范数是等价的，

推论 2 在欧几里得空间内，不同范数意义下的收敛是相互一致的，即在任何一个范数的意义下是收敛的，则在其他范数的意义下也是收敛的。

习题

1. 在 2 维欧几里得空间 R^2 上，是否可以给出范数 $\|\ \|$，使得在这个范数的意义下单位开球

$$O(0,1)=\{x\in R^2 \mid \|x\|<1\}$$

分别由图 1-4 中的 4 个图形表示。

2. 对平面 R^2 上任意一点 $x=(x_1, x_2)$，定义范数

$$\|x\|=\max(|x_1|, |x_2|),$$

由这一范数确定了距离。问：以 $(0,0)$、$(1,0)$、$(0,1)$ 为顶点的三角形是否是等边三角形？在欧几里得范数下怎样？

3. 设 $\|\ \|$ 是 R^n 上的一个范数，

$$O(0,\delta)=\{x\in R^n \mid \|x\|<\delta\}$$

是以原点 O 为中心，半径是 δ 的开球，试证明

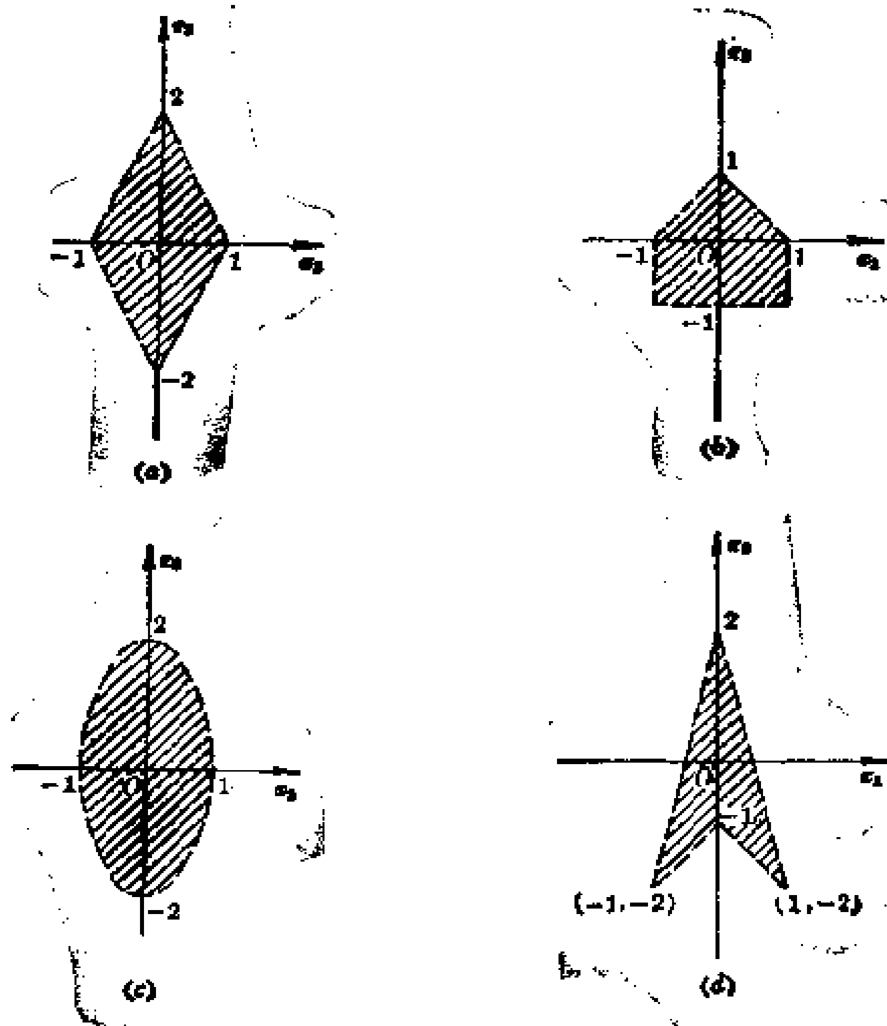


图 1-4

(i) $O(0, \delta)$ 关于原点 O 对称，即对任何 $x\in O(0, \delta)$，有 $-x\in O(0, \delta)$;

(ii) $O(0, \delta)$ 是一个凸集，即对 $O(0, \delta)$ 内任何两点 x 和 y，任何实数 $t\in[0, 1]$，有 $tx+(1-t)y\in O(0, \delta)$，也就是说 x 和 y 的连线在 $O(0, \delta)$ 内.

4. 在 R^n 内，对每一点 $x=(x_1, \cdots, x_n)$ 定义

$$\|x\|_p=(x_1^p+\cdots+x_n^p)^{1/p}, \quad p\geqslant 1,$$

试证明 $\|\ \|_p$ 是 R^n 上的一个范数(特别当 $p=2$ 时就是欧几里得范数).

5. 设 $C^k[a, b]$ 是所有在 $[a, b]$ 上具有 k 阶连续导数的函数组成的线性空间，带有通常的函数的"加"和"数乘". 在 $C^k[a, b]$ 上装备以下三个范数：

$$\|f\|=\max\left\{|f(x)|+\sum_{i=1}^{k}|f^{(i)}(x)|\,\middle|\,a\leqslant x\leqslant b\right\};$$

$$\|f\|' = \int_a^b |f(x)|dx + \sum_{i=1}^{k} \int_a^b |f^{(i)}(x)|dx;$$

$$\|f\|'' = \max\{|f(x)| \mid a \leqslant x \leqslant b\}.$$

问：这三个范数之间有什么关系？

§1.2 欧几里得空间上的线性变换

本节的中心问题是：假设在两个线性空间X和Y之间有一个线性变换T，它将X中的向量x变换为Y中的向量$T(x)$，如果这两个线性空间上各自装备了范数，那么$x\in X$的范数和$T(x)\in Y$的范数之间有什么关系呢？一般说来，它们没有什么必然的联系，但在欧几里得空间上，变换后，$T(x)$的范数却被原先的x的范数所控制，这就是本节的中心论题。

线性变换

设X和Y是两个线性空间，它们可以相同也可以不同，T是从X到Y的一个映射，如果它满足

(i) $T(x+x') = T(x)+T(x')$，　$\forall x, x' \in X$；

(ii) $T(\alpha x) = \alpha T(x)$，　$\forall x \in X$，$\forall \alpha \in R$。

就称T是从X到Y的一个线性变换（或线性映射，线性算子）。

例 1　平面上向量的旋转，拉伸（压缩）都是线性变换，设T：$R^2 \to R^2$是旋转变换，向量v经变换后旋转了一个角度θ（图 1-5），这一变换可以用矩阵

$$T = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

来表示，设$v=(x, y)$，$v'=T(v)=(x', y')$，则$v'=T(v)$，即

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta\cdot x - \sin\theta\cdot y \\ \sin\theta\cdot x + \cos\theta\cdot y \end{pmatrix}.$$

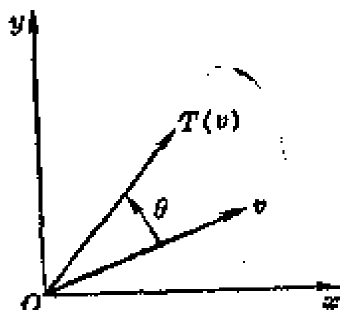


图 1-5

拉伸(或压缩)变换 T_1 也可以用矩阵

$$T_1=\begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}\quad (\alpha>0\text{ 是一个实数})$$

来表示，这时 $v'=T_1(v)$，即

$$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}\alpha x\\ \alpha y\end{pmatrix}.$$

当 $\alpha>1$ 时，是向量的拉伸；当 $0<\alpha<1$ 时，是向量的压缩。

容易验证 T 和 T_1 都是从 R^2 到 R^2 的线性变换。

例 2 R^n 中向量的投影是从 R^n 到 R^n 的一个线性变换。具体的说，设

$$\Pi_i: R^n\to R^n$$

$$(x_1, \cdots, x_n)\mapsto(0, \cdots, 0, x_i, 0, \cdots, 0)$$

它是向量 $x=(x_1, \cdots, x_n)\in R^n$ 在 x_i-轴上的投影(图 1-6)，显然，它是从 R^n 到 R^n 的一个线性变换。

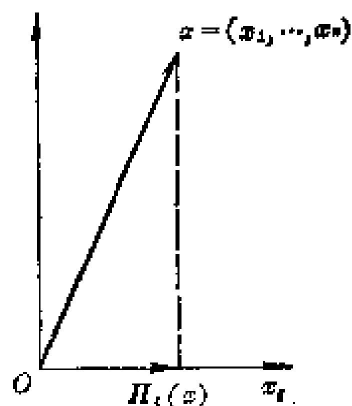


图 1-6

例 3 给定一个 $m\times n$ 矩阵

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix},$$

定义变换 T 如下：$T: R^n\to R^m$，对任何 $x=(x_1, \cdots, x_n)\in R^n$，

$$T(x)=Ax=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \cdots\\ x_n\end{pmatrix},$$

$T(x)$ 是 R^m 中的一个向量。容易验证 T 是从 R^n 到 R^m 的一个线性变换，或者说给定一个 $m\times n$ 矩阵就确定了从 R^n 到 R^m 的一个线性变换。

反过来，在代数学中已经知道，从 R^n 到 R^m 的每一个线性变换，当给定了 R^n 和 R^m 中的基以后，它就可以用一个 $m\times n$ 矩阵来

表示。这一事实将在以后的课文中起重要作用。

例 4 设 $C[a, b]$ 是 $[a, b]$ 上所有连续函数组成的线性空间。定义变换 I 是

$$I: C[a, b] \to R,$$

$$f \mapsto \int_a^b f(x)dx.$$

即在 I 的作用下将 $C[a, b]$ 中的每一个 f 变换为一个实数

$$\int_a^b f(x)dx.$$

由

$$\int_a^b (\alpha f(x) + \beta g(x))dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx$$

知道，I 是从 $C[a, b]$ 到 R 的一个线性变换。

例 5 设 $C^1[a, b]$ 是 $[a, b]$ 上所有具有一阶连续导数的函数组成的线性空间，定义变换 D 是

$$D: C^1[a, b] \to C[a, b]$$

$$f \mapsto \frac{d}{dx} f \quad (\text{即 } f').$$

由

$$\frac{d}{dx}(\alpha f(x) + \beta g(x)) = \alpha \frac{d}{dx} f(x) + \beta \frac{d}{dx} g(x)$$

知道，D 是从 $C^1[a, b]$ 到 $C[a, b]$ 的一个线性变换。

在线性变换中，我们要研究有界的线性变换。设 T 是从 X 到 Y 的线性变换，X 和 Y 不仅是两个线性空间，而且还各自装备了范数 $\| \ \|_X$ 和 $\| \ \|_Y$，它们都是赋范线性空间。如果存在常数 $M>0$，使得对每一个 $x \in X$ 都有

$$\|T(x)\|_Y \leqslant M\|x\|_X,$$

就称 T 是从 X 到 Y 的有界线性变换(或有界线性算子，有界线性映射)。直观的说，在变换 T 的作用下，X 中的每一个点 x 的象 $T(x)$ 的范数 $\|T(x)\|_Y$ 总是被 x 的范数 $\|x\|_X$ 所控制。例 1 中的 R^2 装备了欧几里得范数以后，旋转或拉伸变换就是有界线性变换。以旋转变换 T 为例，设 $v=(x, y) \in R^2$，则

$$T(x)=(\cos\theta\cdot x-\sin\theta\cdot y,\ \sin\theta\cdot x+\cos\theta\cdot y),$$

$$|T(x)|=\sqrt{(\cos\theta\cdot x-\sin\theta\cdot y)^2+(\sin\theta\cdot x+\cos\theta\cdot y)^2}$$
$$=\sqrt{x^2+y^2}=|x|.$$

可见T是有界线性变换。例4中的$C[a,\ b]$装备了上确界范数$\|f\|=\sup\{|f(x)|\,|x\in[a,\ b]\}$以后，变换$I$就是一个有界线性变换，这是因为设$f\in C[a,\ b]$，那么

$$|I(f)|=\left|\int_a^b f(x)dx\right|$$
$$\leqslant(b-a)\cdot\sup\{|f(x)|\,|x\in[a,b]\}$$
$$=(b-a)\cdot\|f\|.$$

线性变换的有界性

线性变换是否一定有界呢？回答是：未必有界，考察例5中的求导算子D。不妨设$[a,\ b]=[0,\ \pi]$，D是从$C^1[0,\ \pi]$到$C[0,\ \pi]$的一个线性变换。在$C^1[0,\ \pi]$和$C[0,\ \pi]$上用同一方式定义范数

$$\|f\|=\sup\{|f(x)|\,|0\leqslant x\leqslant\pi\},$$

其中

$$f\in C^1[0,\ \pi]\ 或者\ f\in C[0,\ \pi].$$

$(C^1[0,\ \pi],\ \|\|)$和$(C[0,\ \pi],\ \|\|)$都是赋范线性空间。令

$$f_n(x)=\sin nx\quad(n=1,\ 2,\ \cdots),\quad x\in[0,\ \pi].$$

我们有

$$f_n\in C^1[0,\ \pi],$$
$$\|f_n\|=\sup\{|\sin nx|\,|0\leqslant x\leqslant\pi\}=1.$$
$$D(f_n)=n\cos nx,\ D(f_n)\in C[0,\ \pi],$$
$$\|D(f_n)\|=\sup\{|n\cos nx|\,|0\leqslant x\leqslant\pi\}=n.$$

显然，不存在这样一个常数$M>0$，使得

$$\|D(f)\|\leqslant M\|f\|,\quad\forall f\in C^1[0,\ \pi],$$

即D不是有界的。

然而，欧几里得空间上的任何线性变换都是有界的，这就是下

面的定理.

定理 1(欧几里得空间上线性变换的有界性) 设在 R^n 和 R^m 上分别装备了范数 $\| \ \|_{R^n}$ 和 $\| \ \|_{R^m}$, T 是从 R^n 到 R^m 的任意一个线性变换,则 T 是有界线性变换.

证明 设 R^n 中的一组基是 $e_1, \cdots, e_n$, R^n 中的每一个向量 x 都表示为

$$x = x_1 e_1 + \cdots + x_n e_n,$$

由 T 的线性,

$$T(x) = x_1 T(e_1) + \cdots + x_n T(e_n),$$

所以

$$\|T(x)\|_{R^m} \leqslant |x_1| \|T(e_1)\|_{R^m} + \cdots + |x_n| \|T(e_n)\|_{R^m},$$

令 $c = \max\{\|f(e_1)\|_{R^m}, \cdots, \|f(e_n)\|_{R^m}\}$,

则
$$\|T(x)\|_{R^m} \leqslant c(|x_1| + \cdots + |x_n|),$$

注意到 $|x_1| + \cdots + |x_n|$ 可以作为 x 的一个范数 (参见1.1例4),再由欧几里得空间上范数的等价性,则存在常数 $c' > 0$, 使得 $|x_1| + \cdots + |x_n| \leqslant c' \|x\|_{R^n}$, 这样便得到

$$\|T(x)\|_{R_m} \leqslant cc' \|x\|_{R^n}.$$

证毕

在定理的条件中, T 是从 R^n 到 R^m 的一个线性变换, 但在证明过程中,我们只利用了 R^n 是有限维的这一特性 (即设它的一组基是 $e_1, \cdots, e_n$), 至于 R^m 也是有限维的这一事实在证明中并没有起什么作用, 从这里可以得到怎样的启发和进一步获得怎样的结论?例如是否可以将 R^m 改为任意一个赋范线性空间$(Y, \| \ \|_Y)$ 呢? 请读者考虑.

定理 2 设 T 是从 R^n 到 R^n 的一个线性变换, 在给定的基下 T 对应的矩阵是 A, 又设 A 的行列式 $\det A \neq 0$, 则存在常数 $m > 0$, 使得对一切 $x \in R^n$, 有

$$\|T(x)\| \geqslant m\|x\|.$$

(注意,上述不等式右端的$\| \ \|$是变换 T 的定义域 R^n 上的范数, 左端的$\| \ \|$是 T 的值域所在的 R^n 上的范数,两者可以相同也可以不同.)

证明并不困难，因为 $\det A\neq 0$，所以 A 可逆，即变换 T 可逆，再利用定理 1 便得证明。其详细的证明留给读者（即下面的习题 3）。

定理 1 和定理 2 表明，从 R^n 到 R^n 的线性变换 T 如果存在逆变换的话，则任何向量 $x\in R^n$ 的范数和它的象 $T(x)$ 的范数相互控制，即存在常数 $c>0$ 和 $d>0$，使

$$c\|x\|\leqslant\|T(x)\|\leqslant d\|x\|,\quad \forall x\in R^n.$$

习　题

1. 设 X、Y、Z 都是线性空间，f 是从 X 到 Y 的线性变换，令 $R_f=\{y\in Y\,|\,y=f(x),\ x\in X\}$，$R_f\subset H\subset Y$，$g$ 是从 H 到 Z 的线性变换，证明 gof 是从 X 到 Z 的线性变换，其中

$$gof(x)=g(f(x)),\quad x\in X.$$

如果 f，g 又都是有界的，gof 如何？

2. 从定理 1 的证明过程中试说明定理 1 可以作怎样的拓广？为什么？

3. 证明定理 2。

4. 设 $C_c(R)$ 是实数轴上所有满足下列条件的函数 f 组成的线性空间，带有通常的函数之间的"加"和"数乘"，(i) f 在 R 连续；(ii) $\{x\in R\,|\,f(x)\neq 0\}$ 的闭包是 R 内的紧集(即有界闭集)。在 $C_c(R)$ 上装备上确界范数

$$\|f\|=\sup\{|f(x)|\,|\,x\in R\},\ f\in C_c(R).$$

又设

$$I:C_c(R)\to R,$$

$$f\mapsto\int_{-\infty}^{\infty}f(x)dx.$$

问：(i) I 是否线性？(ii) I 是否有界？

5. 设 T 是从 R^n 到 R 的一个线性变换，试证明存在一个向量 $c\in R^n$，使对任何 $x\in R^n$，有

$$T(x)=\langle c,x\rangle\quad(\langle\,,\,\rangle\text{是内积}).$$

6. 设 $L(R^n,R^m)$ 是从 R^n 到 R^m 的所有线性变换组成的空间，对任何 T，$S\in L(R^n,R^m)$ 以及任何实数 α，定义

$$(T+S)(x)=T(x)+S(x),\quad (\alpha T)(x)=\alpha T(x),\quad x\in R^n,$$

试证明 $L(R^n,R^m)$ 是一个线性空间，并给出它的一组基。[提示：设 R^n 和 R^m 中的基分别是 $e_1,\cdots,e_n$ 和 $\bar e_1,\cdots,\bar e_m$，考察线性变换 $l_{ij}:R^n\to R^m$，

$l_{ij}(e_i)=e_j$, $l_{ij}(e_k)=0$, $i\neq k$, $(i=1, 2, \cdots, n, j=1, 2, \cdots, m)$。证明 $\{l_{ij}\}$ 是 $L(R^n, R^m)$ 的一组基。]

§1.3 欧几里得空间上的连续映射

本书的一个重要内容是研究欧几里得空间上映射的微分学。因此有必要首先对欧几里得空间上的映射作适当的叙述，特别是连续的映射，这就是本节的内容。

映射

设 D 是 n 维欧几里得空间 R^n 内的一个子集，f 是从 D 到 R^m 的一个映射，即

$$f: D \to R^m,\ D \subset R^n,$$
$$x \mapsto y.$$

这里自变量 $x \in D$ 是 R^n 中的一个点 $x=(x_1, \cdots, x_n)$，D 是 f 的定义域，x 的象 $y=f(x)$ 是 R^m 中的一个点(向量) $y=(y_1, \cdots, y_m)$。因此，也称 f 是定义在 D 内的向量值函数。y 依赖于 x，这表明 y 的每一个坐标 $y_j (j=1, 2, \cdots, m)$ 都依赖于 $x=(x_1, \cdots, x_n)$，也就是说，y_i 是 $(x_1, \cdots, x_n)$ 的函数：

$$\begin{aligned} y_1 &= f_1(x_1, \cdots, x_n), \\ y_2 &= f_2(x_1, \cdots, x_n), \\ &\cdots\cdots\cdots\cdots, \\ y_m &= f_m(x_1, \cdots, x_n) \\ &\quad (x_1, \cdots, x_n \in D), \end{aligned}$$

称 f_i 是 f 的 i-坐标函数，并记

$$f=(f_1, f_2, \cdots, f_m).$$

例 1 设

$$c: [a, b] \to R^3,$$
$$t \mapsto (x, y, z),$$

它是从闭区间 $[a, b]$ 到 R^3 的一个映射。可见坐标 x, y, z 都是 t

的函数，即

$$x=x(t),\quad y=y(t),\quad z=z(t),\quad (a\leqslant t\leqslant b).$$

又如果 x, y, z 都在 $[a, b]$ 连续，我们就说 c 是从 $[a, b]$ 到 R^3 的一个连续映射，它表示 R^3 中的一条连续的曲线。

例 2 设 D 是 R^2 中的一个区域，又设

$$S: D\to R^3$$

$$(u, v)\mapsto(x, y, z).$$

这表明 x, y, z 都是 (u, v) 的函数，即

$$x=x(u, v),\quad y=y(u, v),\quad z=z(u, v),\quad (u, v)\in D.$$

又如果 x, y, z 关于 u、v 有连续的偏导数，并且雅可比行列式

$$\frac{\partial(x,y)}{\partial(u, v)}=\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}\neq 0,\quad (u, v)\in D$$

则这一映射就表示 R^3 中的一个光滑曲面。

例 3 设 D 是 R^3 中的一个区域，又设

$$T: R^3\to R^3$$

$$(u, v, w)\mapsto(x, y, z),$$

于是 x, y, z 分别是 (u, v, w) 的函数：

$$x=f(u, v, w),\quad y=g(u, v, w),$$

$$z=h(u, v, w),\quad (u, v, w)\in D.$$

假设 f, g, h 的具体表达式是

$$x=f(u, v, w)=u\sin v\cdot\cos w,$$

$$y=g(u, v, w)=u\sin v\cdot\sin w,$$

$$z=h(u, v, w)=u\cos v.$$

其中 $(u, v, w)\in D$, $D=\{(u, v, w)\,|\,0<u<+\infty,\ 0<v<\pi,\ 0<w<2\pi\}$，这时映射 T 就是大家熟知的球面坐标变换。

例 4 设 D 是 R^n 中的一个子集，

$$f: D\to R,$$

$$(x_1, \cdots, x_n)\mapsto u,$$

即

$$u=f(x_1, \cdots, x_n),\quad (x_1, \cdots, x_n)\in D.$$

它就是一个 n 元的实值函数。

这几个例子说明了：当引进欧几里得空间上映射的概念以后，就能够统一处理通常数学分析中特别是多元微积分学中的许多重要概念，如曲线、曲面、坐标变换、多元函数等，使我们有可能站在一个更高的层次上来看待这些具体的概念，同时也有可能用一个统一的观点来研究这些各不相同的对象。

设 $f:D\to R^m$，其中 $D\subset R^n$。如果对 D 内任何两个不同的点 x 和 y，其象 $f(x)\neq f(y)$，就称 f 是从 D 到 R^m 的一个单射，也称 f 是一个一一映射。

设 $f:D\to E$，$D\subset R^n$，$E\subset R^m$，如果对 E 内的任何一点 y，在 D 内至少存在一个 x（可能不止一个），使得 $f(x)=y$，或者说 f 的值域 $\{y\in E\,|\,y=f(x),\ x\in D\}=E$，就称 f 是从 D 到 E 的一个满射，它表示函数 f 的值充满了 E，有时也称 f 是从 D 到 E 上的映射。

设 f 是从 D 到 E 的一个满射，同时又是一个单射，就称 f 是从 D 到 E 的一个双射。又称 D 和 E 之间是一一对应的，如果 $f:D\to E$ 是双射，则逆映射 $f^{-1}:E\to D$ 一定存在。

连续映射

设 D 是 R^n 中的一个开集，f 是从 D 到 R^m 的一个映射，$x_0\in D$，如果对任意给定的 $\varepsilon>0$，存在 $\delta>0$，对任何 $x\in D$，当 $|x-x_0|<\delta$ 时，有

$$|f(x)-f(x_0)|<\varepsilon,$$

就称映射 f 在点 x_0 连续。要注意的是：在定义中 $x-x_0$ 的范数 $|\ \ |$ 是 R^n 上的欧几里得范数，而 $f(x)-f(x_0)$ 的范数 $|\ \ |$ 是 R^m 上的欧几里得范数。由欧几里得空间上范数的等价性知道，f 在点 x_0 的连续性与欧几里得空间上范数的选取无关。确切地说：设 R^n 和 R^m 上各自装备了范数 $\|\ \ \|$ 和 $\|\ \ \|'$，D 是 R^n 内的一个开集，$f:D\to R^m$，x_0 是 D 内的一点．如果对任意给定的 $\varepsilon>0$，存在 $\delta>0$，当 $x\in D$ 并且 $\|x-x_0\|<\delta$ 时，有

$$\|f(x)-f(x_0)\|'<\varepsilon,$$

就称 f 在点 x_0 连续（在范数 $\| \ \|$ 和 $\| \ \|'$ 的意义下），欧几里得空间上范数的等价性定理说明了：如果在某个范数的意义下 f 在点 x_0 连续，那么在其他范数的意义下 f 也在点 x_0 连续。为了方便起见，我们常常取欧几里得范数。

f 在点 x_0 连续的概念也可以这样叙述：对任意给定的 $\varepsilon>0$，存在 $\delta>0$，使得

$$f(x)\in O(f(x_0),\ \varepsilon),\qquad \forall x\in O(x_0,\ \delta)\cap D.$$

这里开球 $O(x_0,\ \delta)$ 和开球 $O(f(x_0),\ \varepsilon)$ 分别是由 R^n 和 R^m 上的范数产生的。

粗看，还以为上述定义是一元实值函数在点 x_0 连续的定义，它们在形式上是完全一样的，只不过在现在的情形下 x 和 x_0 是 n 维欧几里得空间内的点，$f(x)$ 和 $f(x_0)$ 是 m 维欧几里得空间内的点，并且各自带有范数罢了。这正是统一处理给我们带来的方便和好处。

如果映射 f 在 D 内每一点连续，就称 f 在 D 内连续，又称 f 是从 D 到 R^m 的一个连续映射。

设 $f=(f_1,\ \cdots,\ f_m)$，f 在点 x_0 连续和它的每一个坐标函数 f_i 在点 x_0 连续有什么关系呢？

定理 1 设点 x_0 是开集 $D\subset R^n$ 内的一点，f 是从 D 到 R^m 的一个映射，$f=(f_1,\ \cdots,\ f_m)$，则 f 在 x_0 连续当且仅当每一个 f_i $(i=1,\ 2,\ \cdots,\ m)$ 在 x_0 连续。

证明 利用不等式

$$|f(x)-f(x_0)|\leqslant\sum_{i=1}^{m}|f_i(x)-f_i(x_0)|,$$

便得到当每个 f_i 在 x_0 连续时，f 也在 x_0 连续。

再利用不等式

$$|f_i(x)-f_i(x_0)|\leqslant|f(x)-f(x_0)|\quad(i=1,\ 2,\ \cdots,\ m)$$

又得到当 f 在 x_0 连续时，每个 f_i 也在 x_0 连续。证毕。

下面的定理 2 是用“开集的语言”来表达连续的概念的。这一

表达方式在形式上已不受范数或距离的束缚，只要有开集的概念就可以用开集来陈述何谓连续。读者可能会想："在一个赋范线性空间中，开集是由范数产生的，追根到底还是范数在起作用。" 确实，在赋范线性空间中是如此，但在更一般的拓扑空间中并无范数而只有开集，一切只能从开集出发，这种用"开集的语言"来定义连续的方式自然可以拓广到一般的拓扑空间上（在第4章中要讨论这一问题）。

定理 2 设 $f:D\to R^m$，D 是 R^n 内的一个开集，那么 f 在 D 连续的必要和充分条件是：对 R^m 内的任一开集 V，它的逆象

$$U=f^{-1}(V)=\{x\in D\mid f(x)\in V\}$$

是 R^n 的一个开集。

在证明之前，我们先用一元实值函数的例子来说明一下。例如实数轴上的正弦函数 $y=f(x)=\sin x$，它是整个实轴上的连续函数，对 y 轴上的开区间 V，它的逆象是 x 轴上开区间 U_1，U_2，…，U_{-1}，U_{-2}，…之并，仍旧是 x 轴上的一个开集(图 1-7)。

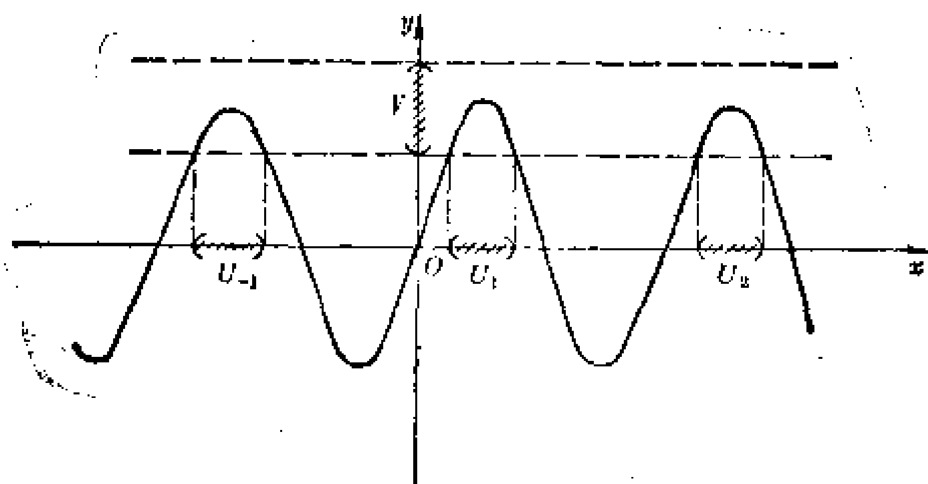


图 1-7

又如图 1-8 中由 $y=f(x)$ 表示的函数，它定义在区间 (a, b) 内，在 (a, b) 内不连续，对 y 轴上的一个开区间 V，它的逆象 $f^{-1}(V)$ 却是 x 轴的一个左闭右开的区间，不是开集。

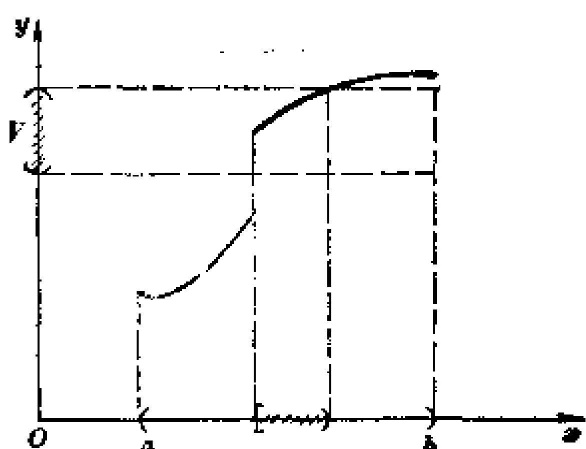


图 1-8

现在证明定理 2。

证明 先证明必要性：设 V 是 R^m 内的任一开集，为了证明其逆象 $f^{-1}(V)$ 是 R^n 中的一个开集，只要证明对每一点 $x_0\in f^{-1}(V)$，一定存在一个开球 $O(x_0,\ \delta)\subset f^{-1}(V)$。

设 $y_0=f(x_0)$，因为 $y_0\in V$，V 是开集，所以存在一个以 y_0 为中心的开球 $O(y_0,\ \varepsilon)\subset V$（图 1-9），再由 f 的连续性，存在 $\delta>0$，当 $x\in O(x_0,\ \delta)$ 时，有

$$f(x)\in O(y_0,\ \varepsilon)\subset V,$$

这表明 $O(x_0,\ \delta)$ 中任何一点 x 的象 $f(x)$ 在 V 中，换句话说，V 的逆象 $f^{-1}(V)$ 包含了 $O(x_0,\ \delta)$，这正是我们所要证明的事。

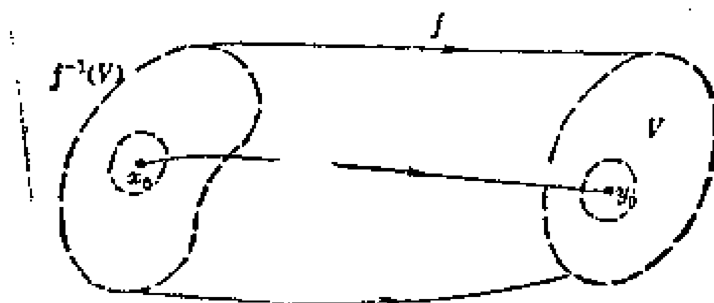


图 1-9

再证明充分性：设 x_0 是 D 内任意一点，其象是 $y_0=f(x_0)$。对任意给定的 $\varepsilon>0$，$V=O(y_0,\varepsilon)$ 是 R^m 中的一个开球（图 1-10），由假设 $f^{-1}(V)$ 是 R^n 中的一个开集（但未必是开球），并且 $x_0\in f^{-1}(V)$。因为 $f^{-1}(V)$ 是开集，所以存在开球 $O(x_0,\ \delta)\subset f^{-1}(V)$，这说明当

$x\in \mathring{O}(x_0, \delta)$时，有 $f(x)\in V$，也就是说，f 在 x_0 连续，而 x_0 是 D 内任一点，这样便证明了充分性。证毕。

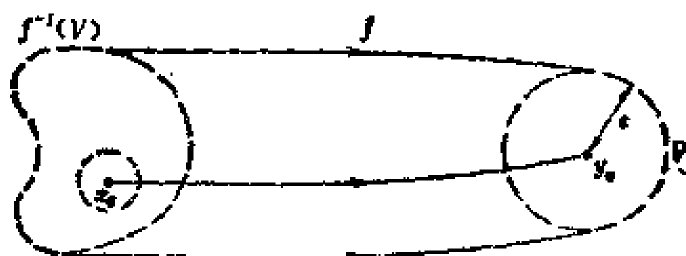


图 1-10

定理 2 说明了连续等价于"开集之逆象仍是开集"，但它是不是又等价于"开集之象仍为开集"呢？或者更明确的问：在连续映射之下，开集的象是否也是开集？回答是否定的。例如实数轴上的连续函数 $y=f(x)=\sin x$，对 x 轴上的开区间 $(0, \pi)$，其象 $f((0, \pi))=\{y\in Y \mid y=\sin x, 0<x<\pi\}=(0, 1]$ 却不是 y 轴上的开集。

下面的定理牵涉到紧集的概念。让我们回忆一下：设 S 是 R^n 内的一个子集，又设 $\{\sigma_\lambda\}_{\lambda\in\Lambda}$ 是 R^n 内一族开集，如果满足

$$\bigcup_{\lambda\in\Lambda}\sigma_\lambda\supset S,$$

（或者说对 S 内的任何一点 x，总存在某个 σ_λ，使 $x\in\sigma_\lambda$），称$\{\sigma_\lambda\}$是 S 的一个开覆盖。如果在 S 的任何一个开覆盖$\{\sigma_\lambda\}$ 中总可以选取出有限个开集

$$\sigma_{\lambda_1}, \sigma_{\lambda_2}, \cdots, \sigma_{\lambda_m}$$

覆盖 S，即 $\bigcup_{i=1}^{m}\sigma_{\lambda_i}\supset S$，则称 S 是 R^n 内的一个紧集。欧几里得空间内的紧集等价于有界闭集，这一结论以后常用到。下面的定理表明在连续映射下，紧集的象仍旧是紧集，这是连续映射的一个重要特性。

定理 3 设 f 是从开集 $D\subset R^n$ 到 R^m 的一个连续映射，又设 $K\subset D$，K 是 R^n 中的一个紧集，则K的象

$$f(K)=\{y\in R^m \mid y=f(x), x\in K\}$$

是 R^m 中的紧集。

证明 设 $\{V_\alpha\}_{\alpha\in\Lambda}$是 R^m 中的一族开集，$\{V_\alpha\}$覆盖了 $f(K)$，要证明的是在 $\{V_\alpha\}_{\alpha\in\Lambda}$ 中存在有限个开集覆盖 $f(K)$。

由定理 2，对每一个开集 V_α，其逆象 $U_\alpha=f^{-1}(V_\alpha)$ 是 R^n 中的一个开集，当 α 遍历 Λ 时，得到 R^n 中一族开集 $\{U_\alpha\}_{\alpha\in\Lambda}$，它们覆盖了 K。因为K是紧集，所以存在有限个

$$U_1=f^{-1}(V_1),\ \cdots,\ U_k=f^{-1}(V_k)$$

覆盖 K，于是 $\{V_\alpha\}_{\alpha\in\Lambda}$ 中的有限个开集

$$V_1,\ V_2,\ \cdots,\ V_k$$

覆盖了 $f(K)$，证毕。

习 题

1. 试证明从 R^n 到 R^m 的线性映射一定是一个连续映射。

2. 设$D\subset R^n$，f 是从D到 R^m 的连续映射，K 是D内一个紧集，试证明 $\{|f(x)|\,\big|\,x\in K\}$ 有最大值和最小值。

3. 设D是 R^n 内的一个开集，$x_0\in D$，f 是从D到 R^m 的连续映射，试证明 f 在点 x_0 连续的必要和充分条件是对 R^m 中任何含 $f(x_0)$ 的开集V，其逆象 $f^{-1}(V)$ 是 R^n 中含 x_0 的开集。

4. 设D是 R^n 中的一个开集，f 是从D到 R^m 的连续映射。又设H是 R^m 中的一个开集，$H\supset\{y\in R^m|y=f(x),\ x\in D\}$，$g$ 是从H到 R^k 的连续映射，试证明$g\circ f$是从D到 R^k 的连续映射。这里 $g\circ f(x)=g(f(x))$，$x\in D$。

5. 如果映射 $f: R^n\to R^m$ 满足条件：R^m 中任一紧集 K，其逆象 $f^{-1}(K)$ 是 R^n 中的紧集。问 f 是连续映射吗？

6. 设映射 $f: R^n\to R^m$，试证明 f 在 R^n 连续的必要和充分条件是：对 R^n 中的任一子集 E，有

$$f(\overline{E})\subset\overline{f(E)}.$$

其中$\overline{E}$是E的闭包；$\overline{f(E)}$是E的象 $f(E)$的闭包。

第2章 可微映射

§2.1 可微映射的概念和性质

偏导数、方向导数、梯度、雅可比（Jacobi）矩阵无疑是数学分析中最重要的基本概念之一，整个多元函数的微分学就是建筑在这些概念的基础上的，众所周知，在这些不同的概念之间存在着非常密切的联系。然而，当我们说这些概念不过是另一个更基本的概念在不同场合下的某种具体表现而已，这或许会令人感到有点新鲜和有趣，同时也会进一步提出“那个更基本的概念”是什么？它有哪些性质？这就是本节要叙述的内容。

引进这样一个“更基本的概念”并不单纯是为了概括和统一处理多元微分学中的若干经典理论，同时还为了将微分学的研究拓广到更一般（例如无限维的）空间上去，具有承上启下的意义。不过后者已超出本书范围。

二元实值函数可微性的拓广

设D是R^2中的一个开集，点$p_0=(x_0, y_0)\in D$，又设函数$f:D\to R$，它是定义在D内的一个二元实数函数，如果它在点p_0可微，按可微的定义，对任何点$p=(x, y)\in D$，有

$$f(x, y)-f(x_0, y_0)=A(p_0)(x-x_0)+B(p_0)(y-y_0)+r(p, p_0),$$

其中$A(p_0)$和$B(p_0)$是仅与点p_0有关而与点p无关的常数，r满足

$$\lim_{|p-p_0|\to 0}\frac{|r(p, p_0)|}{|p-p_0|}=0.$$

容易推断$A(p_0)$和$B(p_0)$分别是偏导数$f_x(x_0, y_0)$和$f_y(x_0, y_0)$。

但上面的形式不便拓广。现在，将它改写成下面的形式：

$$f(x, y)-f(x_0, y_0)=(A(p_0) \quad B(p_0))\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}+r(p, p_0),$$

其中 $(A(p_0)\ B(p_0))$ 是一个 1×2 矩阵，把它看作是从 R^2 到 R 的线性变换，记为 l_{p_0}，$\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$ 是 R^2 中的一点，它就是 $p-p_0=(x-x_0, y-y_0)$，于是上式又可以写为

$$f(p)-f(p_0)=l_{p_0}(p-p_0)+r(p, p_0),$$

$l_{p_0}(p-p_0)$ 是线性变换 l_{p_0} 作用在 $p-p_0$ 上。这一形式和一元实值函数的可微概念多么相似！

仔细观察这一形式就会发现，$p-p_0$ 并不一定非得限制在 R^2 中不可，它完全可以拓广为 R^n（任意正整数 n）中的点，f 的取值也不一定必须限制在 R 内，它也可以拓广为在 R^m（任意正整数 m）中取值，这时，线性变换 l_{p_0} 将成为从 R^n 到 R^m 的线性变换，这一改写的形式却非同小可，由此，读者可以看出一条拓广的路子。

可微

设 D 是 R^n 中的一个开集，x_0 是 D 内的一点。又设 f 是从 D 到 R^m 的一个映射，如果存在一个与 x_0 有关的、从 R^n 到 R^m 的线性变换 l_{x_0}，使得对 D 内的任何点 x，有

$$f(x)-f(x_0)=l_{x_0}(x-x_0)+r(x, x_0),$$

并且 $r(x, x_0)$ 满足下列要求：

$$\lim_{|x-x_0|\to 0}\frac{|r(x, x_0)|}{|x-x_0|}=0,$$

则称映射 f 在点 x_0 可微，称线性变换 l_{x_0} 是 f 在 x_0 的导数（或微分），记为 $Df(x_0)$ 或 $f'(x_0)$。

如果映射 f 在 D 内的每一点可微，就称 f 在 D 可微，又称 f 是 D 内的可微映射。

要注意的是：(i) 在可微定义中，$x-x_0$ 是 R^n 中的点，$f(x)-f(x_0)$ 是 R^m 中的点，$l_{x_0}(x-x_0)$ 是线性变换 l_{x_0} 作用在 $x-x_0$ 上，将

$x-x_0$ 变换为 R^m 中的一点，因此，$r(x, x_0)$ 也是 R^m 中的点。

(ii) 在极限式中，$|x-x_0|\to 0$ 中的 $|\ |$ 是 R^n 上的欧几里得范数，$|r(x, x_0)|$ 中的 $|\ |$ 是 R^m 上的欧几里得范数。由欧几里得空间上范数的等价性可知，可微与范数的选取无关。

(iii) 导数 $Df(x_0)$（或者记为 $f'(x_0)$）不是一个数，而是从 R^n 到 R^m 的线性变换，它与点 x_0 有关，它作用在点 $y\in R^n$ 上，记为 $Df(x_0)(y)$，$f(x_0)(y)$ 是 R^m 内的一点。不要把 $Df(x_0)$ 当作线性变换作用在 x_0 上，$Df(x_0)$ 本身是一个与 x_0 有关的线性变换。用这一观点来看一元实值函数 f，f 在点 x_0 可微是指

$$f(x)-f(x_0)=f'(x_0)(x-x_0)+o(x-x_0),$$

$f'(x_0)$ 是一个数，但应该作如下的理解：它是一个 1×1 矩阵，是从 R 到 R 的一个线性变换，它将任何 $y\in R$ 变换为 $f'(x_0)(y)=f'(x_0)\cdot y$，正好就是数 $f'(x_0)$ 乘 y。

(iv) 可微的概念是和线性化的概念联系在一起的，当我们略去 $r(x, x_0)$（其范数对 $|x-x_0|$ 而言是一个高阶无穷小量）以后，映射 f 就相当于一个常量映射和一个线性映射之和，即在点 x_0 的一个充分小的邻域内，

$$f(x)\approx f(x_0)+Df(x_0)(x-x_0),$$

这一思想在微分学的应用中起重要作用。

可微映射的基本性质

在下面的讨论中，如果不加说明的话，都假定 D 是 R^n 中的一个开集，$x_0\in D$，映射 $f: D\to R^m$。

性质 1 若映射 f 在点 x_0 可微，则它的导数 $Df(x_0)$ 是唯一的。

在一元实值函数的情形下，导数 $f'(x_0)$ 是一种特定形式的极限，再由极限的唯一性，这一结论是不证自明的。但在这里导数不是用极限来定义的，因此无法利用极限的唯一性，而必须另作证明。

证明 由 f 在 x_0 可微，因此对任何 $y\in R^n$，$y\neq 0$，以及任何充分小的实数 λ，有

$$f(x_0+\lambda y)-f(x_0)=l(\lambda y)+\varphi(\lambda y, x_0),$$

l 是从 R^n 到 R^m 的一个线性变换，φ 满足

$$\lim_{\lambda\to 0}\frac{|\varphi(\lambda y, x_0)|}{|\lambda y|}=0。$$

又设 $f(x_0+\lambda y)-f(x_0)$ 还可以另外表示为

$$f(x_0+\lambda y)-f(x_0)=m(\lambda y)+\psi(\lambda y, x_0),$$

m 是从 R^n 到 R^m 的一个线性变换，ψ 满足

$$\lim_{\lambda\to 0}\frac{|\psi(\lambda y, x_0)|}{|\lambda y|}=0。$$

同一个 $f(x_0+\lambda y)-f(x_0)$ 有两个表示式，所以

$$l(\lambda y)-m(\lambda y)=\psi(\lambda y, x_0)-\varphi(\lambda y, x_0),$$

于是

$$\lim_{\lambda\to 0}\frac{|l(\lambda y)-m(\lambda y)|}{|\lambda y|}=\lim_{\lambda\to 0}\frac{|\psi(\lambda y, x_0)-\varphi(\lambda y, x_0)|}{|\lambda y|}=0,$$

再利用 l 和 m 的线性，得

$$\frac{|\lambda|\cdot|l(y)-m(y)|}{|\lambda|\cdot|y|}=\frac{|l(y)-m(y)|}{|y|}\to 0 \quad (\lambda\to 0),$$

所以对任何非 0 的 $y\in R^n$ 有下式成立：

$$l(y)=m(y),$$

而当 $y=0$ 时上式显然成立(因为 l 和 m 都是线性变换)，这样便证明了 $l=m$。证毕。

性质 2 若映射 f 在点 x_0 可微，则 f 在 x_0 连续。反之，则不然。

证明 由 f 在 x_0 可微，得

$$f(x)-f(x_0)=Df(x_0)(x-x_0)+r(x, x_0),$$

$Df(x_0)$ 是从 R^n 到 R^m 的线性变换，所以它是有界的，即对任何 $y\in R^n$，$|Df(x_0)(y)|\leqslant M|y|$，$M$ 是一个常数，则

$$|f(x)-f(x_0)|\leqslant M|x-x_0|+|r(x, x_0)|,$$

再注意到 $|r(x, x_0)|=o(|x-x_0|)$，这样便证明了结论。证毕。

反之，则不然，很容易举出一元实值函数的例子来说明这一结论。

性质 3 设映射 f 和 g 都在点 x_0 可微，α、β 是两个实数，则 $\alpha f+\beta g$ 也在 x_0 可微，并且

$$D(\alpha f+\beta g)(x_0)=\alpha Df(x_0)+\beta Dg(x_0).$$

其证明是明显的。

性质 4（复合映射的链式规则） 设映射 $f:D\to R^m$，D 是 R^n 中的一个开集，点 $a\in D$，f 在 a 可微。又设映射 $g:H\to R^k$，$H\supset f(D)=\{y\in R^m|y=f(x),\ x\in D\}$，$g$ 在 $b=f(a)$ 可微（图 2-1），则复合映射 gof 在 a 可微，并且

$$D(gof)(a)=Dg(b)oDf(a).$$

右端是两个线性变换 $Df(a)$ 和 $Dg(b)$ 的复合，即对任何 $x\in R^n$，

$$Dg(b)oDf(a)(x)=Dg(b)(Df(a)(x)).$$

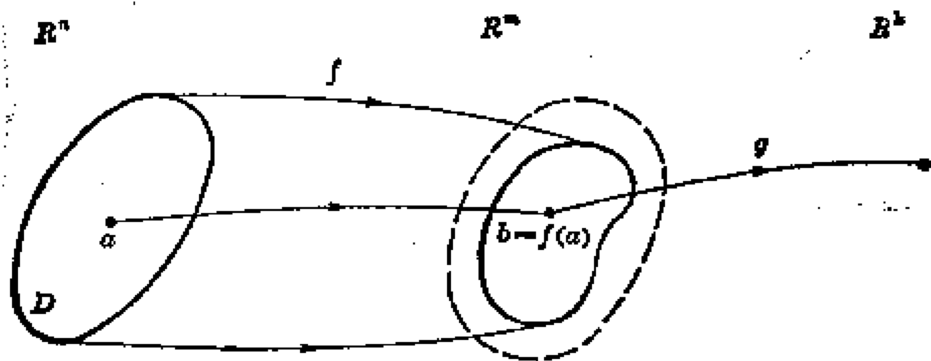


图 2-1

在证明之前，先对这一性质说明一下，在一元函数的微分学中，由两个可微函数复合而成的复合函数的导数是两个函数各自导数的乘积，这个事实是令人费解的，其实，性质 4 告诉我们，复合映射的导数是各自导数的复合，而在一元函数的情形下正好就是各自导数的乘积。

证明 为了使记号简单一些，记 $l=Df(a)$，$m=Dg(b)$，由 f 在 a 可微和 g 在 b 可微，得

$$f(x)-f(a)=l(x-a)+\varphi(x,a),$$

$$\lim_{|x-a|\to 0}\frac{|\varphi(x,a)|}{|x-a|}=0, \tag{1}$$

$$g(y)-g(b)=m(y-b)+\psi(y,b),$$

$$\lim_{|y-b|\to 0}\frac{|\psi(y,\ b)|}{|y-b|}=0. \tag{2}$$

现在考察

$$gof(x)-gof(a)=mol(x-a)+\rho(x,\ a),$$

要证明的是

$$\lim_{|x-a|\to 0}\frac{|\rho(x,\ a)|}{|x-a|}=0.$$

证明如下：由(1)和m是线性变换，得

$$\begin{aligned}\rho(x,\ a)&=g(f(x))-g(f(a))-m(l(x-a))\\&=g(f(x))-g(f(a))-m(f(x)-f(a)-\varphi(x,\ a))\\&=g(f(x))-g(f(a))-m(f(x)-f(a))+m(\varphi(x,a)),\end{aligned}$$

其中$y=f(x)$，$b=f(a)$，所以又有

$$\rho(x,\ a)=g(y)-g(b)-m(y-b)+m(\varphi(x,\ a)),$$

再由(2)，

$$\rho(x,\ a)=\psi(y,\ b)+m(\varphi(x,\ a)).$$

问题化为只要证明下面两个极限成立：

$$\lim_{|x-a|\to 0}\frac{|\psi(y,\ b)|}{|x-a|}=0; \tag{A}$$

$$\lim_{|x-a|\to 0}\frac{|m(\varphi(x,\ a))|}{|x-a|}=0. \tag{B}$$

对任意给定的$\varepsilon>0$，由(2)，存在$\delta>0$，当$|y-b|<\delta$时，有

$$|\psi(y,\ b)|<\varepsilon|y-b|=\varepsilon|f(x)-f(a)|.$$

此外，f在a可微，所以它在a连续，故存在$\eta_1>0$，当$|x-a|<\eta_1$时，$|f(x)-f(a)|<\delta$，于是当$|x-a|<\eta_1$时，上式也成立。再由(2)和线性变换l的有界性，存在$\eta_2>0$，$\eta_2<\eta_1$，当$|x-a|<\eta_2$时，有

$$\begin{aligned}|\psi(y,\ b)|&<\varepsilon|l(x-a)+\varphi(x,\ a)|\\&\leqslant\varepsilon(M|x-a|+\varepsilon|x-a|),\end{aligned}$$

$M>0$是常数。这样便证明了(A)成立。

证明(B)成立是很容易的，由线性变换m的有界性，存在常数$c>0$，使

$$|m(\varphi(x, a))| \leqslant c|\varphi(x, a)|,$$

再由(1)便证明了(B)成立。证毕。

性质 5 设 f 是从 R^n 到 R^m 的常量映射，即对任何 $x \in R^n$，$f(x) = c$（c 是 R^m 中的一个常量），则 f 在 R^n 内可微，并且对任何 $x \in R^n$，

$$Df(x) = 0,$$

右端的 0 是 0 变换，它将 R^n 内的每一点 x 都变换为 R^m 内的 0。

证明是明显的，从略。

在一元函数微分学中，线性变换 $f: R \to R$ 就是"数乘"，即 f 作用在 $x \in R$ 上就是某个确定的数（设它是 a）乘 x，$f(x) = ax$，或者说，f 就是"a 乘"，它的导数 $f'(x) = a$，看作从 R 到 R 的线性变换，它就是 1×1 矩阵，对任何 $y \in R$，$f'(x)(y) = ay$，这表明 f 是"a 乘"，它的导数也是"a 乘"，即 $Df = f$。将这一结论拓广到高维欧几里得空间，就得到下面的性质 6。

性质 6 设 f 是从 R^n 到 R^m 的一个线性变换，则 f 在 R^n 内可微，并且对任何 $x \in R^n$，有

$$Df(x) = f.$$

证明 对每一个 $x_0 \in R^n$，由 f 的线性，得

$$f(x) - f(x_0) = f(x - x_0) + 0,$$

所以 f 在 x_0 可微，并且 $Df(x_0) = f$。证毕。

性质 7 设映射 f 和 g 都是从 D 到 R 的映射（即它们都是定义在 D 内的 n 元实值函数），在点 $x_0 \in D$ 可微，则 $f \cdot g$ 和 f/g（若 $g(x_0) \neq 0$）都在 x_0 可微，并且

$$D(f \cdot g)(x_0) = g(x_0)Df(x_0) + f(x_0)Dg(x_0),$$

$$D\left(\frac{f}{g}\right)(x_0) = \frac{g(x_0)Df(x_0) - f(x_0)Dg(x_0)}{g^2(x_0)}.$$

证明 仅证明乘积：由 f 和 g 在 x_0 的可微性，得

$$f(x) - f(x_0) = Df(x_0)(x - x_0) + \varphi(x, x_0),$$

$$\lim_{|x - x_0| \to 0} \frac{|\varphi(x, x_0)|}{|x - x_0|} = 0,$$

$$g(x) - g(x_0) = Dg(x_0)(x - x_0) + \psi(x, x_0),$$

$$\lim_{|x-x_0|\to 0}\frac{|\psi(x, x_0)|}{|x-x_0|}=0,$$

所以

$$\begin{aligned}f(x)g(x)-f(x_0)g(x_0)&=g(x)(f(x)-f(x_0))\\&\quad+f(x_0)(g(x)-g(x_0))\\&=g(x)(Df(x_0)(x-x_0)+\varphi(x, x_0))\\&\quad+f(x_0)(Dg(x_0)(x-x_0)+\psi(x, x_0)),\end{aligned}$$

g 在 x_0 可微，从而 g 在 x_0 连续，于是在 x_0 的某个邻域内，$g(x)=g(x_0)+\varepsilon$，其中当 $|x-x_0|\to 0$ 时 $\varepsilon\to 0$。代入上式，得

$$\begin{aligned}&f(x)g(x)-f(x_0)g(x_0)\\&\quad=(g(x_0)+\varepsilon)(Df(x_0)(x-x_0)+\varphi(x, x_0))\\&\qquad+f(x_0)(Dg(x_0)(x-x_0)+\psi(x, x_0))\\&\quad=(g(x_0)Df(x_0)+f(x_0)Dg(x_0))(x-x_0)\\&\qquad+r(x, x_0),\end{aligned}$$

$$r(x, x_0)=g(x_0)\varphi+f(x_0)\psi+\varepsilon(Df(x_0)(x-x_0)+\varphi),$$

注意到 $g(x_0)Df(x_0)+f(x_0)Dg(x_0)$ 是从 R^n 到 R 的一个线性变换；$|\varphi|$ 和 $|\psi|$ 是关于 $|x-x_0|$ 的高阶无穷小量；并且 $|Df(x_0)(x-x_0)|\leqslant M|x-x_0|$，$M>0$ 是一个常数，这样便证明了

$$\lim_{|x-x_0|\to 0}\frac{|r(x, x_0)|}{|x-x_0|}=0$$

和

$$D(f\cdot g)(x_0)=g(x_0)Df(x_0)+f(x_0)Dg(x_0).$$ 证毕

性质 8 设 f 和 g 都是从 D(D 是 R 中的一个开集) 到 R^m 的可微映射，定义

$$\langle f, g\rangle(x)=\langle f(x), g(x)\rangle\quad(x\in D).$$

其中 $\langle\, ,\,\rangle$ 是 R^m 上的内积，那么 $\langle f, g\rangle$(它是从 D 到 R 的一个映射) 在任何点 $x\in R$ 可微，并且

$$D\langle f, g\rangle(x)=\langle Df(x)^T, g(x)\rangle+\langle f(x), Dg(x)^T\rangle.$$

$Df(x)^T$ 和 $Dg(x)^T$ 分别是 $Df(x)$ 和 $Dg(x)$ 的转置，$Df(x)$ 和 $Dg(x)$ 都是 $m\times 1$ 矩阵，所以 $Df(x)^T$ 和 $Dg(x)^T$ 都是 $1\times m$ 矩阵，

并把它们看作 R^m 中的点。

其证明留给读者作为习题。

对一个具体的映射 f，如何判断它在某点是否可微，如果它可微，其导数 $Df(x)$ 是什么？从可微的定义出发来作此判断是很困难的，即使对一个结构简单的函数也不是一件容易的事，这是因为必须事前猜测出线性变换 $Df(x)$ 是什么，然后才能估计 $|r(y,x)|=|f(y)-f(x)-Df(x)(y-x)|$ 是不是关于 $|y-x|$ 的高阶无穷小量。在下一节中将给出可微的一个充要条件，利用这个条件将能够比较方便地判断 f 是否可微以及其导数是什么。

习　　题

1. 证明性质 7 中的 f/g 在点 x_0 可微，并且

$$D\left(\frac{f}{g}\right)=\frac{g(x_0)Df(x_0)-f(x_0)Dg(x_0)}{g^2(x_0)}\quad (\text{其中 } g(x_0)\neq 0).$$

2. 证明性质 8(不妨设 $m=2$)。

3. 设可微映射 $f:R^n\to R^n$ 有可微的逆映射 f^{-1}，试证明对任何 $x\in R^n$，有

$$Df^{-1}(y)=(Df(x))^{-1},$$

其中 $y=f(x)$。即逆映射的导数是该映射的导数的逆。

4. 设 $f:R^2\to R$，$f(x,y)=xy$，试证明 f 在任何点 $(a,b)\in R^2$ 可微，并且 $Df(a,b)=(b\ a)$，即对任何 $(x,y)\in R^2$，

$$Df(a,b)(x,y)=(b\ a)\begin{pmatrix}x\\y\end{pmatrix}=bx+ay.$$

5. 设 $f:R^2\to R$，如果满足

$$f(ax,y)=f(x,ay)=af(x,y),\quad (a\text{ 是实数}),$$
$$f(x_1+x_2,y)=f(x_1,y)+f(x_2,y),$$
$$f(x,y_1+y_2)=f(x,y_1)+f(x,y_2),$$

就称 f 是双线性的。试证明：

(i) $$\lim_{\substack{x\to 0\\ y\to 0}}\frac{|f(x,y)|}{\sqrt{x^2+y^2}}=0;$$

(ii) f 在任何点 $(a,b)\in R^2$ 可微，并且对任何 $(x,y)\in R^2$，

$$Df(a,b)(x,y)=f(a,y)+f(x,b).$$

(习题4是它的特殊情形.)

6. 设 $f:R\to R^2$, $f=(f_1,f_2)$, 试证明 f 在 $x_0\in R$ 可微当且仅当 f_1 和 f_2 都在 x_0 可微. 并且这时

$$Df(x_0)=\begin{pmatrix}f_1'(x_0)\\ f_2'(x_0)\end{pmatrix}.$$

7. 设 $l:R^2\to R$, $l(x,y)=x+y$. 又设 f, $g:R^n\to R$, f, g 都在 R^n 可微. 再设 $H:R^n\to R^2$, $H=(f,g)$. 利用 $f+g=l\circ H$, 证明 $f+g$ 可微, 并且

$$D(f+g)(x)=Df(x)+Dg(x).$$

8. 设 f、$g:R^n\to R$, f 和 g 都可微, 利用习题7的思路以及习题4, 证明 $f\cdot g$ 可微, 并且

$$D(fg)(x)=g(x)Df(x)+f(x)Dg(x).$$

§2.2 导数的表示

本节的中心内容是: 当映射 f 在点 x 可微时, 它的导数 $Df(x)$ 是从 R^n 到 R^m 的一个线性变换, 因此, 在 R^n 和 R^m 内给定了一组基以后, $Df(x)$ 可以用一个 $m\times n$ 矩阵表示出来, 但它是怎样的矩阵呢? 这就是本节的中心内容. 同时, 我们还将给出导数 $Df(x)$ 和方向导数、偏导数、梯度向量之间的关系, 使我们有可能用更高的观点来统一处理多元函数微分学中的经典理论.

方向导数、偏导数和梯度向量

设 e_1, e_2, …, e_n 是 R^n 内的一组标准正交基, D 是 R^n 内的一个开集, 点 $x_0\in D$. 又设映射 $f:D\to R$, 它是 n 元实值函数. 如果 f 在 x_0 可微, $Df(x_0)$ 就是从 R^n 到 R 的线性变换, 这个变换作用在 $y\in R^n$ 上是有意义的, 它将 y 变换为一个实数 $Df(x)(y)$, 这个实数含有怎样的意义呢?

定理 1 f 和 x_0 的假设如同上面所说的, 则对任何 $y\in R^n$, $|y|=1$, 有

$$Df(x_0)(y)=\lim_{t\to 0}\frac{f(x_0+ty)-f(x_0)}{t}.$$

等式的右端对我们说来并不陌生，它是函数 f 在点 x_0 沿方向 y 的方向导数。定理 1 表明导数 $Df(x_0)$ 作用在单位向量 y 上就得到沿方向 y 的方向导数。

证明 因为 f 在 x_0 可微，所以

$$f(x_0+ty)-f(x_0)=Df(x_0)(ty)+r(ty,\ x_0),$$

$$\lim_{t\to 0}\frac{|r(ty,\ x_0)|}{|ty|}=0.$$

再由 $Df(x_0)$的线性

$$Df(x_0)(ty)=tDf(x_0)(y)$$

和 $|y|=1$，使得到

$$\lim_{t\to 0}\frac{f(x_0+ty)-f(x_0)}{t}=Df(x_0)(y)+\lim_{t\to 0}\frac{r(ty,\ x_0)}{|ty|}$$

$$=Df(x_0)(y).$$

证毕

从定理的证明中可以看出，如果 $y\in R^n$，$y\neq 0$ 但 $|y|\neq 1$，则仍有

$$\lim_{t\to 0}\frac{f(x_0+ty)-f(x_0)}{t}=Df(x_0)(y)$$

成立，不过它不是函数 f 在点 x_0 沿单位方向 $\frac{y}{|y|}$ 的方向导数，两者相差一个常数因子。

推论 在定理 1 的条件下，

$$Df(x_0)(e_i)=\frac{\partial}{\partial x_i}f(x_0).$$

e_i 是 R^n 内的 i-基向量。这说明导数 $Df(x_0)$ 作用在基向量 e_i 上便得到 f 在点 x_0 的关于 x_i 的偏导数。

证明 在定理 1 中特别取 $y=e_i=(0,\ \cdots,\ 0,\ 1,\ 0,\ \cdots,\ 0)$，并设 $x_0=(x_1^0,\ x_2^0,\ \cdots,\ x_n^0)$，则

$$Df(x_0)(e_i)=\lim_{t\to 0}\frac{f(x_0+te_i)-f(x_0)}{t}$$

$$= \lim_{t \to 0} \frac{f(x_1^0, \cdots, x_i^0 + t, \cdots, x_n^0) - f(x_1^0, \cdots, x_i^0, \cdots, x_n^0)}{t}$$

$$= \frac{\partial}{\partial x_i} f(x_0).$$

证毕

在定理1的条件下，$Df(x_0)$是从R^n到R的线性变换，给定了R^n中的一组基$e_1, \cdots, e_n$以后，$Df(x_0)$可以用一个$1 \times n$矩阵表示，这就是下面的定理2。

定理2 在定理1的假设下

$$Df(x_0) = (f_{x_1}(x_0) \quad f_{x_2}(x_0) \quad \cdots \quad f_{x_n}(x_0)).$$

它是一个$1 \times n$矩阵，其中$f_{x_i}(x_0)$是$\frac{\partial}{\partial x_i} f(x_0)$的另一个写法。

证明 设$e_1, e_2, \cdots, e_n$是R^n的一组标准正交基。又设l是从R^n到R的一个线性变换，我们知道，l将由它作用于基向量$e_i (i = 1, 2, \cdots, n)$上的值所完全确定，即若$l(e_i) = \lambda_i$，则$l$所对应的$1 \times n$矩阵就是

$$(\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n).$$

现在，由定理1的推论得$Df(x_0)$（它是从R^n到R的线性变换）作用于基向量e_i上等于$\frac{\partial f}{\partial x_i}(x_0)$，所以$Df(x_0)$对应的$1 \times n$矩阵就是

$$\left(\frac{\partial f}{\partial x_1}(x_0) \quad \frac{\partial f}{\partial x_2}(x_0) \quad \cdots \quad \frac{\partial f}{\partial x_n}(x_0)\right).$$

证毕

如果我们把$(f_{x_1}(x_0), \cdots, f_{x_n}(x_0))$看作一个向量，则这个向量就是函数$f$（或者说数量场$f$）在点$x_0$的梯度。

推论1 设多元实值函数f在点x_0可微，并且x_0又是f的一个极值点[即存在点x_0的邻域$O(x_0, \delta)$，对任何$x \in O(x_0, \delta)$，都有$f(x) \leqslant f(x_0)$或者都有$f(x) \geqslant f(x_0)$]，则

$$Df(x_0) = 0，或者写成 f'(x_0) = 0.$$

其证明是明显的。

推论1的论述形式完全和一元实值函数时的情形一样，对不同的对象有一个统一的形式处理。

推论 2 设 f 是一个三元实值函数，它在点 x_0 可微，又设 y 是 R^3 中的一个单位向量，方向余弦是 $\cos\alpha$、$\cos\beta$、$\cos\gamma$，则 f 在点 x_0 沿方向 y 的方向导数是

$$Df(x_0)(y)=(f_{x_1}(x_0)\quad f_{x_2}(x_0)\quad f_{x_3}(x_0))\begin{pmatrix}\cos\alpha\\ \cos\beta\\ \cos\gamma\end{pmatrix}$$

$$=f_{x_1}(x_0)\cos\alpha+f_{x_2}(x_0)\cos\beta+f_{x_3}(x_0)\cos\gamma.$$

证明由定理 1 和定理 2 即可得出。

雅可比矩阵

现在将定理 2 的结果加以拓广。设 D 是 R^n 中的一个开集，点 $x_0\in D$，映射 $f:D\to R^m$，　$f=(f_1, \cdots, f_m)$，又设 f 在 x_0 可微，则其导数 $Df(x_0)$ 在 R^n 和 R^m 中给定的标准正交基下，可以用怎样的 $m\times n$ 矩阵来表示呢？

定理 3 设映射 $f:D\to R^m$，$f=(f_1, \cdots, f_m)$，D 是 R^n 中的开集，$x_0\in D$，则 f 在 x_0 可微的必要和充分条件是：每一个 f_i 在 x_0 可微。此外，若 f 在 x_0 可微，则

$$Df(x_0)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0)\\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_n}(x_0)\\ \cdots & \cdots & \cdots & \cdots\\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0)\end{pmatrix}.$$

上式右端的 $m\times n$ 矩阵称为映射 f 的雅可比矩阵，或者称它是函数组 $f_1, \cdots, f_m$ 的雅可比矩阵。当 $m=n$ 时，称雅可比矩阵的行列式为雅可比行列式。

证明 先证明充分性，设每个 $f_i(i=1, 2, \cdots, m)$ 在 x_0 可微，即对 $x\in D$，有

$$f_i(x)-f_i(x_0)=Df_i(x_0)(x-x_0)+r_i(x, x_0),$$

$$\lim_{|x-x_0|\to 0}\frac{|r_i(x, x_0)|}{|x-x_0|}=0 \quad (i=1, 2, \cdots, m).$$

由定理 2 知道

$$Df_i(x_0)=\left(\frac{\partial f_i}{\partial x_1}(x_0) \quad \frac{\partial f_i}{\partial x_2}(x_0) \quad \cdots \quad \frac{\partial f_i}{\partial x_n}(x_0)\right).$$

记

$$J(x_0)=\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix},$$

它表示从 R^n 到 R^m 的一个线性变换，$Df_i(x_0)$ 是 $J(x_0)$ 的第 i 行的行向量。对 $x\in D$，考察

$$f(x)-f(x_0)=J(x_0)(x-x_0)+r(x, x_0),$$

要证明的是

$$\lim_{|x-x_0|\to 0}\frac{|r(x, x_0)|}{|x-x_0|}=0.$$

设 $x_0=(x_1^0, \cdots, x_n^0)$，$x=(x_1, \cdots, x_n)$，因为

$$J(x_0)(x-x_0)=\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}\begin{pmatrix} x_1-x_1^0 \\ \vdots \\ x_n-x_n^0 \end{pmatrix}$$

$$=\begin{pmatrix} Df_1(x_0)(x-x_0) \\ Df_2(x_0)(x-x_0) \\ \cdots\cdots \\ Df_m(x_0)(x-x_0) \end{pmatrix},$$

所以

$$r(x, x_0)=\begin{pmatrix} f_1(x)-f_1(x_0)-Df_1(x_0)(x-x_0) \\ f_2(x)-f_2(x_0)-Df_2(x_0)(x-x_0) \\ \cdots\cdots \\ f_m(x)-f_m(x_0)-Df_m(x_0)(x-x_0) \end{pmatrix},$$

$$\frac{|r(x, x_0)|}{|x-x_0|} \leqslant \sum_{i=1}^{m} \frac{|f_i(x)-f_i(x_0)-Df_i(x_0)(x-x_0)|}{|x-x_0|}$$
$$=\sum_{i=1}^{m} \frac{|r_i(x, x_0)|}{|x-x_0|},$$

上式右端每一项当 $|x-x_0|\to 0$ 时是趋于0的，这样便证明了映射 f 在点 x_0 可微，并且还证明了

$$Df(x_0)=\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}.$$

证明必要性：设映射 f 在 x_0 可微，作投影 $\pi_i(i=1, 2, \cdots, m)$：

$$\pi_i: R^m \to R$$
$$(y_1, \cdots, y_m) \mapsto y_i,$$

容易验证 π_i 是从 R^m 到 R 的线性变换，因此它是在 R^m 内可微的。而 $f_i=\pi_i\circ f$，即对任何 $x\in D$，通过 f 映射为 R^m 中的一点 $f(x)=(f_1(x), \cdots, f_m(x))$，再通过 π_i 映射为 $f_i(x)$。利用复合映射的可微性（2.1 性质 4）便得出 $f_i(1\leqslant i\leqslant m)$ 在点 x_0 可微。

定理的后一部分，其证明已经包含在上面的推断过程中。证毕。

要注意的是，设映射 $f=(f_1, \cdots, f_m)$ 中的每一个 f_i 在点 x_0 的偏导数 $\frac{\partial f_i}{\partial x_1}(x_0), \cdots, \frac{\partial f_i}{\partial x_n}(x_0)$ 都存在，这时虽然可以从形式上写出相应的雅可比矩阵，但还不能说 f 在点 x_0 可微，也不能说这个矩阵是 f 在 x_0 的导数，这是因为仅仅偏导数存在并不足以保证 f_i 在 x_0 可微。例如在 R^2 内，设 $f: R^2\to R$，它的定义是

$$f(x, y)=\begin{cases} \frac{xy}{x^2+y^2}, & (x, y)\neq(0, 0) \\ 0, & (x, y)=(0, 0) \end{cases}$$

在点(0, 0)，f 的两个偏导数都存在并且 $f_x(0,0)=0$，$f_y(0,0)=0$，从形式上可以写出雅可比矩阵 $(0\ 0)$，但 f 在原点不可微(事实上 f 在原点不连续)，因此矩阵 $(0\ 0)$ 不是 f 在原点的导数。但是，如果每一个偏导数 $\frac{\partial f_i}{\partial x_j}$ $(i=1, 2, \cdots, m;\ j=1, 2, \cdots, n)$ 都在点 x_0 连续，则由数学分析知道 f_i 必在 x_0 可微，从而 f 在 x_0 可微，这时又称 f 在点 x_0 连续可微。

设 $f=(f_1, \cdots, f_m)$，每一个 f_i 在开集 $D\subset R^n$ 内的偏导数不仅存在而且连续，就称 f 是从 D 到 R^m 的一个 C^1 类映射，记为 $f\in C^1$。我们再从另一个观点来看这件事，如果 f 在 D 可微，那么对 D 内任何一点 x，线性变换 $Df(x)$ 有意义，换句话说，对每一个 $x\in D$，存在一个线性变换 $Df(x)$ 与此 x 对应，我们把 $Df(x)$ 看作 x 的象，这样便得到一个映射(记它是 Df)，Df 定义在 D 内，取值是(从 R^n 到 R^m 的)线性变换，由于在给定的基下，线性变换可以用一个 $m\times n$ 矩阵来表示，因此映射 Df 是定义在 D 内，取值是 $m\times n$ 矩阵，事实上，对每个 $x\in D$，$Df(x)$ 是一个 $m\times n$ 的雅可比矩阵。

设 M 是所有 $m\times n$ 矩阵组成的赋范线性空间 (见 1.1 例 3)，那么映射 Df 就是

$$Df: D\to M$$

$$x\mapsto f \text{ 在 } x \text{ 的雅可比矩阵。}$$

如果映射 Df 在 D 连续，就称 f 是一个 C^1 类映射。

这里有两个问题需要说明：

(i) 何谓映射 Df 在 D 连续；

(ii) C^1 类映射的两个定义是否等价。

先说明第一个问题：用比较直观的语言来说，所谓连续，就是指对 D 内任何一点 x_0，当 $x\in D$ 并且 x 充分接近 x_0 时，f 在点 x 的雅可比矩阵 $J(x)$ 和在点 x_0 的雅可比矩阵 $J(x_0)$ 也充分接近。确切的说，设 f 在点 x 的雅可比矩阵是 $\left(\frac{\partial f_i}{\partial x_j}(x)\right)_{\substack{i=1,\cdots,m\\ j=1,\cdots,n}}$，此外，

我们在D内采用欧儿里得范数，在M内采用1·1例3中的范数，则“连续”的含意是：对任意给定的$\varepsilon>0$，存在$\delta>0$，当$x\in D$，$|x-x_0|<\delta$时，有

$$\|J(x)-J(x_0)\|=\left\{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{\partial f_i}{\partial x_j}(x)-\frac{\partial f_i}{\partial x_j}(x_0)\right)^2\right\}^{1/2}<\varepsilon.$$

有了这样一个确切的表达以后，就容易回答第二个问题了。

定理 4 C^1类的两个定义是等价的。

证明 设Df是从D到M的连续映射，x_0是D内的任意一点，则对任意给定的$\varepsilon>0$，存在$\delta>0$，当$x\in D$并且$|x-x_0|<\delta$时，有

$$\left|\frac{\partial f_i}{\partial x_j}(x)-\frac{\partial f_i}{\partial x_j}(x_0)\right|$$
$$\leqslant\left\{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{\partial f_i}{\partial x_j}(x)-\frac{\partial f_i}{\partial x_j}(x_0)\right)^2\right\}^{1/2}<\varepsilon,$$

这就证明了$\dfrac{\partial f_i}{\partial x_j}$ $(i=1,2,\cdots,m,\ j=1,2,\cdots,n)$在点$x_0$连续。

反过来，如果每一个$\dfrac{\partial f_i}{\partial x_j}$在$D$连续，$x_0$是$D$内的任意一点，则对任意给定的$\varepsilon>0$，存在$\delta>0$，当$x\in D$，$|x-x_0|<\delta$时，有

$$\left|\frac{\partial f_i}{\partial x_j}(x)-\frac{\partial f_i}{\partial x_j}(x_0)\right|<\frac{\varepsilon}{\sqrt{mn}}\quad(i=1,2,\cdots,m;\ j=1,2,\cdots,n).$$

于是

$$\left\{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{\partial f_i}{\partial x_j}(x)-\frac{\partial f_i}{\partial x_j}(x_0)\right)^2\right\}^{1/2}<\varepsilon,$$

即映射Df在D连续。证毕。

和C^1类映射相仿，我们定义C^2类映射。设$f=(f_1,\cdots,f_m)$，如果每个f_i在开集$D\subset R^n$内有二阶连续偏导数$\dfrac{\partial^2 f_i}{\partial x_j\partial x_k}$ $(i=1,2,\cdots,m;\ j、k=1,2,\cdots,n)$，就称$f$是$D$内的一个$C^2$类映射，记为$f\in C^2$。也可以用下面的观点来看这件事：设$f\in C^1$，则映射$Df$是从$D$到$M$的连续映射，其中$M$是由所有$m\times n$矩阵组成的赋范线性空间。现在，我们要把$M$看作一个欧儿里得空间，看法

是这样的：在M和$n+m$维欧几里得空间R^{n+m}之间建立一个对应关系，设$A=(a_{ij})_{\substack{i=1,2,\cdots,m\\j=1,2,\cdots,n}}$，把$A$和$R^{n+m}$内的点$(a_{11},\cdots,a_{1n},a_{21},\cdots,a_{2n},\cdots,a_{m1},\cdots,a_{mn})$对应起来，这一关系建立了$M$和$R^{n+m}$之间的一一对应。并且有下列特性：

(i) 保持了双方的“加”和“数乘”。即设A、$B\in M$分别和a、$b\in R^{n+m}$对应，则$A+B$必和$a+b$对应，λA必和λa对应(λ是实数)。或者说作为线性空间两者是同构的。

(ii) 范数不变，即设$A\in M$和$a\in R^{n+m}$对应，则

$$\|A\|=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}^2\right)^{1/2}=|a|,$$

或者说作为赋范空间两者是等距的。

由这两个原因，我们可以把M和R^{n+m}看作同一个赋范线性空间，所不同的仅仅是空间中元素的记号而已。作了这样的注解之后，我们可以说映射Df是从D到R^{n+m}的连续映射，如果它在D内可微的话，就称Df的导数$D(Df)$是f在D内的二阶导数(或二阶微分)，记为D^2f。又如果$Df\in C^1$，就称f是一个C^2类映射。同样可以证明C^2类的两个定义是等价的。

一般的，设$f=(f_1,\cdots,f_m)$，如果每个f_i在开集$D\subset R^2$内对各个变元$x_j(j=1,2,\cdots,n)$有k阶连续偏导数，就称f是一个C^k类函数，记为$f\in C^k$。又如果每一个f_i对各个变元有任意阶连续导数，就称f是一个C^∞类映射，记为$f\in C^\infty$。

例1 设映射$f:R^2\to R$，$(x,y)\mapsto z$，定义如下：

$$u=u(x,y),$$

$$v=v(x,y,u),$$

$$w=w(y,v),$$

$$z=z(x,v,w).$$

假定每一个函数都可微。求Df。

从f的定义可知道它是一个复合映射

$$f=f_4\circ f_3\circ f_2\circ f_1,$$

$$(x, y) \xrightarrow{f_1} (x, y, u) \xrightarrow{f_2} (x, y, v) \xrightarrow{f_3} (x, v, w) \xrightarrow{f_4} z,$$

f_1, f_2, f_3, f_4 的定义分别如下：

$$f_1: (x, y) \to (x, y, u), \text{ 其中}$$

$$x = x, \quad y = y, \quad u = u(x, y),$$

f_1 在任何点(x, y)的导数是

$$Df_1(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ u_x(x, y) & u_y(x, y) \end{pmatrix}.$$

$$f_2: (x, y, u) \mapsto (x, y, v), \text{ 其中}$$

$$x = x, \quad y = y, \quad v = v(x, y, u),$$

f_2 在任何点(x, y, u)的导数是

$$Df_2(x, y, u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ v_x(x, y, u) & v_y(x, y, u) & v_u(x, y, u) \end{pmatrix},$$

$$f_3: (x, y, v) \mapsto (x, v, w), \text{ 其中}$$

$$x = x, \quad v = v, \quad w = w(y, v).$$

f_3 在任何点 (x, y, v) 的导数是

$$Df_3(x, y, v) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & w_y(y, v) & w_v(y, v) \end{pmatrix},$$

$$f_4: (x, v, w) \mapsto z, \text{ 其中}$$

$$z = z(x, v, w),$$

f_4 在任何点(x, v, w)的导数是

$$Df_4(x, v, w) = (z_x(x, v, w) \quad z_v(x, v, w) \quad z_w(x, v, w))$$

再由复合映射的链式规则得(为书写简单起见，略去自变量的记号)

$$Df = Df_4 \circ Df_3 \circ Df_2 \circ Df_1$$

$$= (z_x \quad z_v \quad z_w) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & w_y & w_v \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ v_x & v_y & v_u \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ u_x & u_y \end{pmatrix}$$

$$=(f_x \quad f_y),$$

所以

$$f_x = z_x + z_v(v_x + v_u u_x) + z_w(w_v(v_x + v_u u_x)),$$
$$f_y = z_v(v_y + v_u u_y) + z_w(w_y + w_v(v_y + v_u u_y)).$$

例 2 设 $c:(0,1)\to R^3$, $c(t)=(x(t),y(t),z(t))$, 并且 c 在 $(0,1)$ 连续可微，又设对 $(0,1)$ 内的每一个 t, $Dc(t)\neq 0$, 这时 c 表示 R^3 内一条光滑曲线。把 $Dc(t)$ 看作向量

$$Dc(t)=(x'(t),y'(t),z'(t)),$$

它就是曲线 c 在点 $(x(t),y(t),z(t))$ 的切向量。

如果我们在复平面 C 和 R^2 之间建立一个对应关系，每一个复数 $z=x+iy$ 和 R^2 内的点 (x,y) 一一对应起来，任何一个复变函数 $f(x+iy)=u(x,y)+iv(x,y)$ 和从 R^2 到 R^2 的映射 $f=(f_1,f_2)$ 一一对应起来。其中 $f_1(x,y)=u(x,y)$, $f_2(x,y)=v(x,y)$。在这样的对应下，复变函数的复导数和从 R^2 到 R^2 的映射的导数有什么关系呢？

例 3 考察复变函数

$$f(z)=z^2=(x+iy)^2=(x^2-y^2)+i2xy,$$

和它对应的从 R^2 到 R^2 的映射 $f=(f_1,f_2)$ 是

$$f_1(x,y)=x^2-y^2, \quad f_2(x,y)=2xy.$$

对复变函数 f 来说，它是可微的，

$$f'(z)=2z=2(x+iy).$$

所以对任何复数 $h=h_1+ih_2$,

$$f'(z)\cdot h=2[(xh_1-yh_2)+i(xh_2+yh_1)].$$

对映射 $f=(f_1,f_2)$ 来说，也是可微的，

$$Df(x,y)=\begin{pmatrix}\frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y}\\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y}\end{pmatrix}=2\begin{pmatrix}x & -y\\ y & x\end{pmatrix},$$

所以对 R^2 内的任何点 $h=(h_1,h_2)$, $Df(x,y)$ 作用在 (h_1,h_2) 上，得

$$Df(x,y)(h)=2\begin{pmatrix}x & -y\\ y & x\end{pmatrix}\begin{pmatrix}h_1\\ h_2\end{pmatrix}$$

$$= 2(xh_1 - yh_2, yh_1 + xh_2),$$

它所对应的复数是 $2[(xh_1 - yh_2) + i(yh_1 + xh_2)]$，这正好就是 $f'(z)\cdot h$. 这表明作为复变函数的 f 和作为从 R^2 到 R^2 的映射的 f，两者都是可微的，并且两者的导数是一致的。

但一般说来，未必如此。

例 4 考察复变函数

$$f(z) = \bar{z} = \overline{(x+iy)} = x - iy.$$

由于它的实部和虚部不满足柯西-黎曼(Cauchy-Riemann)方程，所以 f 是不可微的. 但和它对应的映射 $f = (f_1, f_2)$ 是

$$f_1(x, y) = x, \quad f_2(x, y) = -y,$$

显然，f 在任何点 (x, y) 是可微的，并且

$$Df(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

如果 f 是一个复的解析函数，那么把它看作复变函数，或者把它看作从 R^2 到 R^2 的映射，两者都可微，并且导数是一致的。其证明留给读者，作为习题.

在这一节中给出了导数 $Df(x)$ 的矩阵表示，利用矩阵表示也可以证明上一节中可微映射的一些基本性质，证明过程可能比上一节中的证明更方便些，读者不妨试试看。那么为什么要在上一节中加以证明呢？在上一节的证明中，我们只利用 $Df(x)$ 是一个有界线性变换这一特性，根本不涉及 $Df(x)$ 的矩阵表示，这样做至少会启发我们把可微的概念和可微映射的基本性质拓广到更一般的赋范线性空间(例如巴拿哈(Banach)空间)上去，即使在那里有界的线性变换不能用一个 $m\times n$ 矩阵表示，但只要有有界线性变换就可以用相仿的方法给出可微的定义和可微映射的基本性质.

习　题

1. 设映射 $f: R^2 \to R^4$, $(u, v) \mapsto (w, x, y, z)$，它的定义是

$$w = u + v, \quad x = u^2 + v^2, \quad y = 4uv + u^2v^2, \quad z = (u^2+1)^2 - 4v,$$

求 $Df(0,0)$，$Df(1,2)$。

2. 设映射 $f:R^3\to R^3$，$f=(f_1,f_2,f_3)$。

$$u=f_1(x,y,z)=xyz,\quad v=f_2(x,y,z)=(x+y)^2-2z,$$
$$w=f_3(x,y,z)=e^{x-2y}.$$

求 $Df(1,1,1)$ 和 $Df(x,y,z)$。

3. 设 $f:R^4\to R^2$，

$$f(x,y,u,v)=(xy(\cos u+\sin v),\ xe^{u+v}-y),$$

求 $Df\left(1,2,0,\frac{\pi}{2}\right)$、$Df(0,-1,\pi,\pi)$。

4. 设 $g:R^2\to R$，$(x,y)\mapsto w$，它的定义是

$$u=u(x,y),$$
$$v=v(x,y,u),$$
$$w=w(x,u,v).$$

其中 u，v，w 都是可微函数。将 g 分解为若干个映射的复合，求 g 的导数 Dg。

5. 设 f 是复平面 C 上的解析函数，将它看成 $R^2\to R^2$ 的映射，证明后者在 R^2 可微，并且两者的导数是一致的。

6. 利用导数的矩阵表示，证明 2-1 节中的性质 7 和性质 8。

§2.3 逆映射定理

"可逆性"问题是数学中一类非常重要的问题，例如一个矩阵是否存在逆阵；一个变换是否存在逆变换；一个映射（或算子）是否存在逆映射（或逆算子）等等，在数学的理论和应用上都具有非常重要的价值。本节的中心是研究欧几里得空间上可微映射在什么条件下存在可微的逆映射，并利用它来研究隐函数的存在性问题。

逆映射定理

我们将采取和一元实值函数对比的方式来叙述本节的中心定理，这样做的目的是希望能够让读者看到如何对照一元实值函数的结果来猜测多元向量值函数的相应的结果，然后再去论证它是

否正确：

中心定理(逆映射定理)

一元实值函数

设 $f:(c, d)\to R$，

f 在 (c, d) 连续可微，

又设 $a\in(c, d)$，$f'(a)\neq 0$，

则存在含 a 的开区间 U 和含 $b=f(a)$ 的开区间 V，使

(i) $f:U\to V$ 是双射；

(ii) f^{-1} 在 V 连续可微，且

$f^{-1\prime}(y)=(f'(x))^{-1}$，

这里 $y=f(x)$，$x\in U$。

多元向量值函数

设 $f:D\to R^n$，D 是 R^n 内的开集，f 在 D 连续可微，

又设 $a\in D$，$\det Df(a)\neq 0$，

则存在含 a 的开集 U 和含 $b=f(a)$ 的开集 V，使

(i) $f:U\to V$ 是双射；

(ii) f^{-1} 在 V 连续可微，且

$Df^{-1}(y)=(Df(x))^{-1}$，

这里 $y=f(x)$，$x\in U$。

这个定理说明了，在定理的条件下，逆映射 f^{-1} 在一个局部的范围内是存在的，且是连续可微的，并且逆映射的导数是原来映射的导数的逆。

在证明中心定理之前，我们先证明以下几个定理。这几个定理并不是专门为本节的中心定理而设置的，它们各自有其自身的价值，在其他场合下也能起作用。

定理 1 设映射 $f:D\to R^m$，$D\subset R^n$ 是开集。又设 f 在 D 连续可微，则对任何 $x, y\in D$，

$$f(y)-f(x)=Df(x)(y-x)+r(y, x),$$

在任何紧集 $K\subset D$ 上一致成立

$$\lim_{|y-x|\to 0}\frac{|r(y, x)|}{|y-x|}=0.$$

注：如果 f 仅仅在 D 可微，则对每个固定的 $x\in D$，总有

$$f(y)-f(x)=Df(x)(y-x)+r(y, x)\quad (y\in D),$$

$$\lim_{|y-x|\to 0}\frac{|r(y, x)|}{|y-x|}=0,$$

上面的极限用 ε-δ 的语言表达出来，就是对任意给定的 $\varepsilon>0$，存在 $\delta>0$，这个 δ 依赖于 ε 和 x，当 $y\in D$ 并且 $|y-x|<\delta$ 时，有

$$\frac{|r(y, x)|}{|y-x|}<\varepsilon.$$

定理1说明了：如果在映射 f 上再加条件，f 在 D 连续可微，则在任何紧集 $K\subset D$ 上，上述极限是一致成立的，即 δ 仅与 ε 有关，而与 $x\in K$ 无关。

证明 设 $f=(f_1, \cdots, f_m)$，$x=(x_1, \cdots, x_n)$，$y=(y_1, \cdots, y_n)$，则

$$Df(x)(y-x)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x)\\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x)\end{pmatrix}\begin{pmatrix}y_1-x_1\\ \vdots\\ y_n-x_n\end{pmatrix}$$

$$=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x)(y_1-x_1)+\cdots+\frac{\partial f_1}{\partial x_n}(x)(y_n-x_n)\\ \cdots\cdots\cdots\cdots\\ \frac{\partial f_m}{\partial x_1}(x)(y_1-x_1)+\cdots+\frac{\partial f_m}{\partial x_n}(x)(y_n-x_n)\end{pmatrix}.$$

再设 $r(y, x)=(r_1(y, x), \cdots, r_m(y, x))$，其中

$$r_i(y, x)=f_i(y)-f_i(x)-\Big(\frac{\partial f_i}{\partial x_1}(x)(y_1-x_1)+\cdots+\frac{\partial f_i}{\partial x_n}(x)(y_n-x_n)\Big).$$

利用微分学的中值定理

$$f_i(y)-f_i(x)=\frac{\partial f_i}{\partial x_1}(\xi_i)(y_1-x_1)+\cdots+\frac{\partial f_i}{\partial x_n}(\xi_i)(y_n-x_n),$$

ξ_i 在 x 和 y 的连线上（由于 D 是开集，因此当 x, y 充分接近时，x 和 y 的连线必在 D 内），将上式代入 r_i 中，得

$$r_i(y, x)=\Big(\frac{\partial f_i}{\partial x_1}(\xi_i)-\frac{\partial f_i}{\partial x_1}(x)\Big)(y_1-x_1)+\cdots+\Big(\frac{\partial f_i}{\partial x_n}(\xi_i)-\frac{\partial f_i}{\partial x_n}(x)\Big)(y_n-x_n),$$

于是

$$|r|\leqslant\sum_{i=1}^{n}|r_i|\leqslant\sum_{i=1}^{m}\sum_{j=1}^{n}\left|\frac{\partial f_i}{\partial x_j}(\xi_i)-\frac{\partial f_i}{\partial x_j}(x)\right||y_j-x_j|.$$

由假设 f 在 D 连续可微，即每一个 $\frac{\partial f_i}{\partial x_j}(i=1, 2, \cdots, m, j=1, 2, \cdots, n)$ 在 D 连续，所以 $\frac{\partial f_i}{\partial x_j}$ 在紧集 K 上是一致连续的，因此对任意给定的 $\varepsilon>0$，存在与 ε 有关而与 $x\in K$ 无关的 $\delta>0$，当 x、$y\in K$，并且 $|y-x|<\delta$ 时，有

$$\left|\frac{\partial f_i}{\partial x_j}(\xi_i)-\frac{\partial f_i}{\partial x_j}(x)\right|<\varepsilon \quad (i=1, 2, \cdots, m; j=1, 2, \cdots, n).$$

再注意到 $|y_j-x_j|\leqslant|y-x|\quad(j=1, 2, \cdots, n)$，这样便得到

$$|r|\leqslant n^2\varepsilon|y-x|.$$

证毕

定理 2 设映射 $f:D\to R^m$，$D\subset R^n$ 是开集，又设 f 在 D 连续可微，$a\in D$，则对任意给定的 $\varepsilon>0$，存在开球 $O(a, \delta)$，使得当 $x\in O(a, \delta)$ 时，对一切 $y\in R^n$，有

$$|Df(x)(y)-Df(a)(y)|<\varepsilon|y|.$$

定理 2 表明在连续可微的条件下，在点 a 的近傍线性变换 $Df(x)$ 可以用线性变换 $Df(a)$ 近似代替。

证明 设 $y=(y_1, \cdots, y_n)$，因为

$$Df(x)(y)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x)\\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x)\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},$$

$$Df(a)(y)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a)\\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a)\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},$$

所以

$$Df(x)(y)-Df(a)(y)$$

$$=\begin{pmatrix}\left(\frac{\partial f_1}{\partial x_1}(x)-\frac{\partial f_1}{\partial x_1}(a)\right)y_1+\cdots+\left(\frac{\partial f_1}{\partial x_n}(x)-\frac{\partial f_1}{\partial x_n}(a)\right)y_n\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \left(\frac{\partial f_m}{\partial x_1}(x)-\frac{\partial f_m}{\partial x_1}(a)\right)y_1+\cdots+\left(\frac{\partial f_m}{\partial x_n}(x)-\frac{\partial f_m}{\partial x_n}(a)\right)y_n\end{pmatrix},$$

$$|Df(x)(y)-Df(a)(y)|\leqslant\sum_{i=1}^{m}\sum_{j=1}^{n}\left|\frac{\partial f_i}{\partial x_j}(x)-\frac{\partial f_i}{\partial x_j}(a)\right||y_j|.$$

利用 $\frac{\partial f_i}{\partial x_j}$ $(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 在点 a 的连续性，即对任意给定的 $\varepsilon>0$，存在 $\delta>0$，当 $|x-a|<\delta$ 时，

$$\left|\frac{\partial f_i}{\partial x_j}(x)-\frac{\partial f_i}{\partial x_j}(a)\right|<\varepsilon/mn\quad(i=1,2,\cdots,m;j=1,2,\cdots,n),$$

于是

$$|Df(x)(y)-Df(a)(y)|<\varepsilon|y|.$$

证毕

在定理 2 的结论中，开球 $O(a,\delta)$ 可以改为闭球 $\overline{O}(a,\eta)$，即当 $x\in\overline{O}(a,\eta)$ 时，对一切 $y\in R^n$，有

$$|Df(x)(y)-Df(a)(y)|<\varepsilon|y|.$$

这是因为只要取 $\eta<\delta$ 就可以了。

定理 3 在中心定理的条件下，存在闭球 $\overline{O}(a,\eta)$ 和常数 $C>0$, $D>0$，使

$$C|y-x|\leqslant|f(y)-f(x)|\leqslant D|y-x|,\quad x、y\in\overline{O}(a,\eta).$$

证明的思想方法是比较简单的，映射 f 的可微性保证了在点 x 的局部范围内有

$$f(y)-f(x)\approx Df(x)(y-x),$$

f 的连续可微性又保证了当点 x 充分接近点 a 时，$Df(x)$ 可以用 $Df(a)$ 近似代替，而 $\det Df(a)\neq0$ 意味着存在常数 $C'>0$ 和 $D'>0$，使得

$$C'|y-x|\leqslant|Df(a)(y-x)|\leqslant D'|y-x|.$$

把这些事实综合起来就得到定理 3。

证明定理 3：

$$\begin{aligned}f(y)-f(x)&=Df(x)(y-x)+r(y,x)\\&=Df(a)(y-x)+Df(x)(y-x)\\&\quad-Df(a)(y-x)+r(y,x),\end{aligned}$$

其中

$$\lim_{|y-x|\to0}\frac{|r(y,x)|}{|y-x|}=0.$$

所以

$$|f(y)-f(x)|\leqslant|Df(a)(y-x)|+|Df(x)(y-x)-Df(a)(y-x)|+|r(y,x)|.$$

注意下列事实：

(i) 在定理 2 中，令 $\varepsilon=1$，则存在闭球 $\overline{O}(a,\eta_1)$，当 $x\in\overline{O}(a,\eta_1)$时，有

$$|Df(x)(y-x)-Df(a)(y-x)|<|y-x|.$$

(ii) 由定理 1，上面的极限在D内任何紧集上是一致的，可以认为 η_1 是适当的小，使得

$$\frac{|r(y,x)|}{|y-x|}<1,\quad x、y\in\overline{O}(a,\eta_1).$$

(iii) 线性变换 $Df(a)$是有界的。即存在常数 $d'>0$，使得

$$|Df(a)(x)|\leqslant d'|x|,\ x\in R^n.$$

这样便得到

$$|f(y)-f(x)|\leqslant d'|y-x|+|y-x|+|y-x|=(d'+2)|y-x|,\quad x、y\in\overline{O}(a,\eta_1).$$

另一方面，又有

$$|f(y)-f(x)|\geqslant|Df(a)(y-x)|-|Df(x)(y-x)-Df(a)(y-x)|-|r(y,x)|,$$

注意下列事实：

(i) 在 $\det Df(a)\neq0$ 的条件下，存在常数 $C'>0$，使 $|Df(a)(x)|\geqslant C'|x|$，$x\in R^n$。

(ii) 存在闭球 $\overline{O}(a,\eta_2)$，当 $x,y\in\overline{O}(a,\eta_2)$时，有

$$|Df(x)(y-x)-Df(a)(y-x)|<\frac{C'}{3}|y-x|,$$

$$|r(y,x)|<\frac{C'}{3}|y-x|.$$

这样便得到

$$|f(y)-f(x)|>\frac{C'}{3}|y-x|,\quad x,y\in\overline{O}(a,\eta_2).$$

令 $\eta=\min(\eta_1,\eta_2)$，$\overline{O}(a,\eta)$即为所求的闭球。证毕。

推论 在中心定理的条件下，存在闭球$\overline{O}(a, \eta)$，使得f是$\overline{O}(a, \eta)$上的单射。

证明 由定理3，存在闭球$\overline{O}(a, \eta)$和常数$C>0$，对任何x，$y\in\overline{O}(a, \eta)$，有

$$|f(y)-f(x)|\geqslant C|y-x|.$$

可见当$x\neq y$时必有$f(x)\neq f(y)$。证毕。

定理4 设F是R^n中的一个闭集，$y_0\notin F$，则存在开球$O(y_0, \delta)$，对$O(y_0, \delta)$内的每一点y和F内的每一点y'，有下式成立（图2-2）：

$$|y-y_0|<|y-y'|.$$

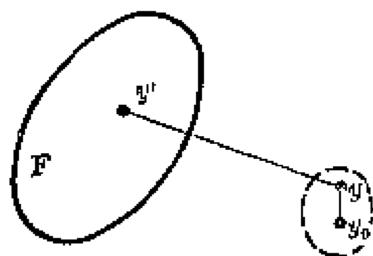


图 2-2

证明 $y_0\notin F$即$y_0\in F^c$（F^c是F的补集R^n-F），F是闭集，所以F^c是开集，于是存在开球$O(y_0,\eta)\subset F^c$，再令$\delta=\eta/3$，$O(y_0,\delta)$即为所求。证毕。

现在证明中心定理：

由定理3，存在闭球$\overline{O}(a,\eta)$，当$x,y\in\overline{O}(a,\eta)$时，有

$$c|y-x|\leqslant|f(y)-f(x)|\leqslant d|x-y|,$$

其中$c>0,d>0$是常数，同时可见f是$\overline{O}(a,\eta)$内的单射。

对每个$x\in D$，$\det Df(x)$是$n\times n$矩阵$Df(x)$的行列式，它的展开式是由$\dfrac{\partial f_i}{\partial x_j}$ $(i=1,2,\cdots,n;\ j=1,2,\cdots,n)$经过有限次的代数运算（和，差，积）组成的，因此$\det Df$是$D$内的连续函数。已知$\det Df(a)\neq 0$，所以存在以点$a$为中心的一个闭球，不妨设它就是$\overline{O}(a,\eta)$，使得

$$\det Df(x)\neq 0,\quad x\in\overline{O}(a, \eta).$$

设$\overline{O}(a, \eta)$的边界是S，它是R^n中的一个球面，是一个紧集，所以它的象$f(S)$也是一个紧集（1.3定理3）。

很明显$a\notin S$，而f是$\overline{O}(a,\eta)$上的单射，所以$b=f(a)\notin f(S)$（图2-3），由定理4，存在以点b为中心的开球V，使

$$|y-f(a)|<|y-f(x)|,\ y\in V,\ x\in S.$$

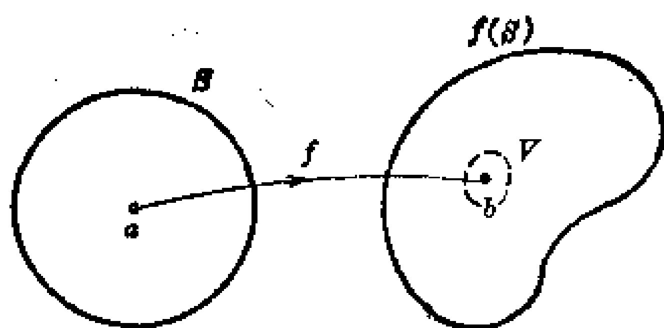


图 2-3

(i) 先证明存在含点 a 的开集 U，使 $f:U\to V$，f 是双射，这就是说 f 不仅是单射，并且对每一个 $y\in V$，存在一个 $x\in U$，使 $y=f(x)$。

由前面的论述，f 显然是一个单射。再设 $y_0=(y_1^0, \cdots, y_n^0)$ 是 V 内任意给定的一点，作映射 $g:\overline{O}(a, \eta)\to R$，它的定义是

$$g(x)=|y_0-f(x)|^2=\sum_{i=1}^{n}(y_i^0-f_i(x))^2,$$

每一个 f_i 在 $\overline{O}(a, \eta)$ 连续，所以 g 在 $\overline{O}(a,\eta)$ 连续，而 $\overline{O}(a,\eta)$ 是紧集，因此 g 在 $\overline{O}(a, \eta)$ 上的最小值存在，设这个最小值是 α，并且 $f(x_0)=\alpha$，　$x_0\in\overline{O}(a, \eta)$。

点 x_0 会不会在边界 S 上呢？对边界 S 上的每一点 x，有 $|y_0-f(a)|<|y_0-f(x)|$，这表明 x_0 不在 S 上，即 x_0 在开球 $O(a, \eta)$ 内。再由极值的必要条件 $Dg(x_0)=0$，这就是

$$\sum_{i=1}^{n}(y_i^0-f_i(x_0))\frac{\partial f_i}{\partial x_j}(x_0)=0 \quad (j=1, 2, \cdots, n).$$

但在 $O(a, \eta)$ 内 $\det Df(x)\neq 0$，所以上面的齐次方程组只有全为 0 的解：

$$y_i^0-f_i(x_0)=0 \quad (i=1, 2, \cdots, n),$$

这就是说，存在 $x_0\in O(a, \eta)$，使 $y_0=f(x_0)$。记

$$U=\{x\in O(a, \eta)\,|\,y=f(x), y\in V\}=f^{-1}(V),$$

由 f 的连续性可知道，U 是一个开集，并且 U 显然含有 a，同时 $f:U\to V$ 是双射，于是 $f^{-1}:V\to U$ 存在。

(ii) 再证明 f^{-1} 在 V 连续：

对 x_1、$x_2 \in U$，设 $y_1 = f(x_1)$、$y_2 = f(x_2)$，或者写为 $x_1 = f^{-1}(y_1)$，$x_2 = f^{-1}(y_2)$。由定理 3，

$$C|x_1 - x_2| \leqslant |f(x_1) - f(x_2)|,$$

也可以改写为

$$C|f^{-1}(y_1) - f^{-1}(y_2)| \leqslant |y_1 - y_2|,$$

由此即得 f^{-1} 在 V 内是连续的。

(iii) 最后证明 f^{-1} 在 V 连续可微，并且

$$Df^{-1}(y) = (Df(x))^{-1},$$

其中 $y = f(x)$，$x \in U$。设 x_0 是 U 内的任意一点，$y_0 = f(x_0)$，由 f 在 x_0 可微，得

$$f(x) - f(x_0) = Df(x_0)(x - x_0) + r(x, x_0),$$

$$\lim_{|x-x_0| \to 0} \frac{|r(x, x_0)|}{|x - x_0|} = 0.$$

在 $\det Df(x_0) \neq 0$ 的条件下，线性变换 $Df(x_0)$ 的逆变换 $(Df(x_0))^{-1}$ 存在，将 $(Df(x_0))^{-1}$ 作用于上面的等式，得

$$(Df(x_0))^{-1}(f(x) - f(x_0)) = x - x_0 + (Df(x_0))^{-1}(r(x, x_0)),$$

又可以写为

$$(Df(x_0))^{-1}(y - y_0) = f^{-1}(y) - f^{-1}(y_0) + (Df(x_0))^{-1}(r(x, x_0)),$$

这就是

$$f^{-1}(y) - f^{-1}(y_0) = (Df(x_0))^{-1}(y - y_0) \\ - (Df(x_0))^{-1}(r(x, x_0)).$$

如果能够证明

$$\lim_{|y-y_0| \to 0} \frac{|(Df(x_0))^{-1}(r(x, x_0))|}{|y - y_0|} = 0,$$

那么就证明了 f^{-1} 在 y_0 可微，并且其导数是 $(Df(x_0))^{-1}$。现在证明这件事：

因为 $(Df(x_0))^{-1}$ 是线性变换，所以存在常数 $L > 0$，使

$$|(Df(x_0))^{-1}(r(x, x_0))| \leqslant L|r(x, x_0)|,$$

问题化为只要证明

$$\lim_{|y-y_0|\to 0} \frac{|r(x,x_0)|}{|y-y_0|}=0.$$

这是不难证明的。利用

$$\frac{|r(x,x_0)|}{|y-y_0|}=\frac{|r(x,x_0)|}{|x-x_0|}\quad\frac{|x-x_0|}{|y-y_0|},$$

注意到

$$|x-x_0|\to 0 \text{ 当且仅当 } |y-y_0|\to 0,$$

$$|x-x_0|=0 \text{ 当且仅当 } |y-y_0|=0,$$

$$\lim_{|x-x_0|\to 0}\frac{|r(x,x_0)|}{|x-x_0|}=0,$$

$$C|x-x_0|\leqslant|y-y_0|.$$

便证明了上面的结论。

为什么说 f^{-1} 在 V 连续可微呢？已经证明了 $Df^{-1}(y)=(Df(x))^{-1}$，而矩阵 $Df(x)$ 中的元素是 $\frac{\partial f_i}{\partial x_j}(x)$，它们都在 U 内连续，再加上 $\det Df(x)\neq 0$，因此其逆阵中的每个元素也都在 U 内连续，即 f^{-1} 在 V 连续可微．证毕。

还要说明两件事：

(i) $\det Df(a)\neq 0$ 并非必要条件。例如一元实值函数 f，$f(x)=x^3$，在 $x=0$ 处 $f'(0)=0$，但反函数 f^{-1} 仍旧是存在的。

(ii) 如果 $\det Df(a)=0$，又如果 f^{-1} 在点 $b=f(a)$ 的某个邻域内存在，那么可以断言 f^{-1} 在点 b 必不可微。这是因为

$$f^{-1}\circ f(x)=x,\quad x\in U,$$

其中 U 是含点 a 的某个邻域，假设 f^{-1} 在点 b 可微，则

$$Df^{-1}(b)\circ Df(a)=\begin{pmatrix}1 & & 0\\ & 1 & \\ & & \ddots \\ 0 & & 1\end{pmatrix},$$

于是

$$\det Df^{-1}(b)\cdot\det Df(a)=1,$$

这和 $\det Df(a)=0$ 相矛盾。

隐函数定理

设 f 是定义在 $R^n\times R^m$ 内的一个映射，在 R^m 内取值。问在什么条件下可以从方程

$$f(x,y)=0 \quad (x\in R^n, y\in R^m)$$

中确定 y 是 x 的函数 $y=g(x)$。用坐标写出来就是：设有函数方程组

$$\begin{cases} f_1(x_1,\cdots,x_n,y_1,\cdots,y_m)=0,\\ f_2(x_1,\cdots,x_n,y_1,\cdots,y_m)=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,\\ f_m(x_1,\cdots,x_n,y_1,\cdots,y_m)=0,\end{cases}$$

问在什么条件下，可以由方程组确定

$$\begin{cases} y_1=g_1(x_1,\cdots,x_n),\\ y_2=g_2(x_1,\cdots,x_n),\\ \cdots\cdots\cdots\cdots\cdots\cdots,\\ y_m=g_m(x_1,\cdots,x_n).\end{cases}$$

这就是隐函数的存在性问题。下面将利用逆映射定理来研究它。叙述方式仍旧采用同数学分析中两个变量的隐函数定理对比的方法来进行：

两个变量的情形

定理 5（隐函数定理）

设 $f: R\times R\to R$

$$(x,y)\mapsto u=f(x,y),$$

如果 f 满足：

(i) 在含点 (a,b) 的某个开集内连续可微。

(ii) $f(a,b)=0$,

(iii) $f_y(a,b)\neq 0$,

一般情形

定理 5（隐函数定理）

设 $f: R^n\times R^m\to R^m$

$$(x,y)\mapsto u=f(x,y),$$

$$x\in R^n, y\in R^m$$

如果 f 满足：

(i) 在含点 (a,b) 的某个开集内连续可微。

(ii) $f(a,b)=0$,

(iii) $\det f_y(a,b)\neq 0$.

这里

$$f_y(a,b) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) \cdots \frac{\partial f_1}{\partial y_m}(a,b) \\ \cdots\cdots\cdots\cdots \\ \frac{\partial f_m}{\partial y_1}(a,b) \cdots \frac{\partial f_m}{\partial y_m}(a,b) \end{pmatrix}.$$

则存在含点 a 的开区间 A 和含点 b 的开区间 B，对每一个 $x \in A$，存在唯一的 $y = g(x) \in B$，满足

$$f(x, g(x)) = 0,$$

并且 g 在 A 连续可微。

则存在含点 a 的开集 A 和含点 b 的开集 B，对每一个 $x \in A$，存在唯一的 $y = g(x) \in B$，满足

$$f(x, g(x)) = 0,$$

并且 g 在 A 连续可微。

下面给出两个变量情形的证明过程，对一般情形的证明可以和它作对比，其证明留给读者，作为习题。

证明 作

$$\varphi: R \times R \to R \times R,$$
$$(x, y) \mapsto (x, u),$$
$$x = x, \quad u = f(x, y).$$

则 φ 连续可微，并且

$$D\varphi(a,b) = \begin{pmatrix} 1 & 0 \\ f_x(a,b) & f_y(a,b) \end{pmatrix}$$

$$\det D\varphi(a,b) = f_y(a,b) \neq 0,$$

由逆映射定理，在含 (a,b) 的某个开集内，存在连续可微的逆映射 $\varphi^{-1}: (x, u) \mapsto (x, y)$，其中

$$x = x, \quad y = h(x, u).$$

令 $u = 0$，记 $y = h(x, 0) = g(x)$，有：g 在含点 a 的某个开区间内连续可微，并且 $0 = f(x, g(x))$，这个 g 就是所要求的。证毕。

给定方程组

$$\begin{cases} f_1(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots, \\ f_m(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0. \end{cases}$$

设它满足隐函数定理的条件，如何求 y_i 关于 x_j 的偏导数呢？或者说，如何求定理中的 g 的导数 Dg 呢？将每一个方程两端对 x_j 求导，得

$$\frac{\partial f_1}{\partial x_j}+\frac{\partial f_1}{\partial y_1}\frac{\partial y_1}{\partial x_j}+\cdots+\frac{\partial f_1}{\partial y_m}\frac{\partial y_m}{\partial x_j}=0,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$\frac{\partial f_m}{\partial x_j}+\frac{\partial f_m}{\partial y_1}\frac{\partial y_1}{\partial x_j}+\cdots+\frac{\partial f_m}{\partial y_m}\frac{\partial y_m}{\partial x_j}=0,$$

由 $\det f_y(x,y)\neq 0$，即上述线性方程组的行列式非 0，从而可以解出

$$\frac{\partial y_1}{\partial x_j},\cdots,\frac{\partial y_m}{\partial x_j}\quad (j=1,2,\cdots,n).$$

雅可比行列式的几何意义

雅可比矩阵和雅可比行列式在映射的微分学中扮演了非常重要的角色，这里再从几何上对雅可比行列式作解释，使我们对它的理解臻于完善。

设 $a_1,a_2,\cdots,a_n$ 是 R^n 中线性无关的 n 个向量，令

$$A=\{x\in R^n \mid x=\alpha_1a_1+\cdots+\alpha_na_n, 0\leqslant\alpha_1,\cdots,\alpha_n\leqslant 1\},$$

称 A 是 R^n 中的一个 n 维闭平行 $2n$ 面体。当 $a_1,a_2,\cdots,a_n$ 两两正交时，称 A 是 n 维闭矩形。

设 $e_1,e_2,\cdots,e_n$ 是 R^n 的一组标准正交基，$a_i=a_{i1}e_1+\cdots+a_{in}e_n$ $(i=1,2,\cdots,n)$，则 A 的容积 $v(A)$ 是

$$v(A)=|\det(a_1,a_2,\cdots,a_n)|,$$

其中 $\det(a_1,a_2,\cdots,a_n)$ 是由向量 $a_1,a_2,\cdots,a_n$ 组成的行列式，即

$$\det(a_1,a_2,\cdots,a_n)=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}.$$

又设 l 是从 R^n 到 R^n 的一个非奇线性变换（即 l 是线性变换，并且它所对应的矩阵的行列式 $\det l\neq 0$），则在 l 的作用下，将 n 维闭矩形 A 变换为 n 维闭平行 $2n$ 面体 $l(A)$，它们的容积之

间有如下关系式：

$$v(l(A)) = |\det l|v(A)。$$

很明显，如果A是n维开矩形，$l(A)$就是n维开平行$2n$面体，上述结论也成立。

这一结论还可以推广到R^n内的有界开集上去。设D是R^n中的一个有界开集，则

$$v(l(D)) = |\det l|v(D)。$$

要作这样的拓广，关键在于如何定义开集D的容积$v(D)$。要定义开集的容积，其依据是下面的引理。

引理(开集的构造) 在欧几里得空间内，任何有界开集D都可以表示为一列n维闭矩形$A_1, A_2, \cdots, A_n, \cdots$的并：$D = \bigcup_{k=1}^{\infty} A_k$，并且这些闭矩形的内部(即除边界外)两两不相交。

这一事实是相当直观的，以平面而论，设D是平面上的一个开集，用平行于坐标轴的直线$x=i, y=j$ $(i, j=0, \pm1, \pm2, \cdots)$组成的网格将整个平面划分为一列闭正方形，这些闭正方形分为两类；第一类是整个正方形都含在D内；第二类是其它的正方形。将第一类闭正方形全部取出来，设它们是$A_{11}, A_{12}, \cdots, A_{1n_1}$，然后将网格加密，用直线$x=i$，$y=j$ $\left(i, j=0, \pm\frac{1}{10}, \pm\frac{2}{10}, \cdots\right)$划分整个平面，但除去已经取出来的那些$A_{11}, \cdots, A_{1n_1}$，同样产生两类闭正方形，再将所有含在$D$内的闭正方形取出来，设它们是$A_{21}, A_{22}, \cdots, A_{2n_2}$。如此继续下去，取出来的那一列闭正方形$A_{11}, \cdots, A_{1n_1}, A_{21} \cdots, A_{2n_2}, A_{31}, \cdots$即为所求。

有了开集的构造，就可以定义有界开集D的容积。设

$$D = \bigcup_{k=1}^{\infty} A_k, A_k \text{是闭矩形，其内部两两不相交.}$$

定义D的容积$v(D)$是

$$v(D) = \sum_{k=1}^{\infty} v(A_k)。$$

当D是有界集时，$v(D) < \infty$。

剩下一个问题是：D的分解显然不是唯一的，设D有两种分解 $D=\bigcup_{k=1}^{\infty}A_k$，$D=\bigcup_{k=1}^{\infty}B_k$，$A_k$,$B_k$ 都是闭矩形，并且 A_k 的内部两两不相交，B_k 的内部也两两不相交，则是否有

$$\sum_{k=1}^{\infty}v(A_k)=\sum_{k=1}^{\infty}v(B_k)$$

成立？回答是肯定的。请读者考虑，为方便起见，可以考虑平面的情形。同时不难发现，如果将D分解为一列内部两两不相交的闭平行多面体 A_k 的并，则也有

$$v(D)=\sum_{k=1}^{\infty}v(A_k).$$

到此，我们所要求的结论

$$v(l(D))=|\det l|\,v(D)$$

是很明显的了。

对线性变换的研究可启发我们讨论非线性的变换，对任意一个可微的映射 f，一般说来，它未必是线性的，但在点 x 的充分小的邻域内，可以把它近似的看作线性映射 $Df(x)$ 和一个常量映射之和，这使我们可猜测到在可微映射 f 的作用下，n 维矩形 A 和它的象 $f(A)$的容积之间岂不是也可能有

$$v(f(A))\approx|\det Df(x)|\,v(A)?$$

确实有这一结论。

定理 6（雅可比行列式的几何解释） 设D是 R^n 内的一个开集，$f:D\to R^n$ 满足：(i) f 在D连续可微；(ii) f 是单射；(iii) 雅可比行列式 $\det Df(x)\neq0(x\in D)$。又设S是D内的 n 维开矩形，它以点 $x=(x_1,\cdots,x_n)\in D$ 为中心，

$$S=\left\{(y_1,\cdots,y_n)\in R\,\Big|\,|y_i-x_i|<\frac{h}{2},i=1,2,\cdots,n\right\},$$

(即 S 的每一个边的边长是 h)，则

$$\lim_{h\to0}\frac{v(f(S))}{v(S)}=|\det Df(x)|.$$

并且这一极限在D内任意一个紧集K上是一致的，也就是说，如

果用 ε-δ 的方式来给出上述极限的定义时，δ 仅与 ε 有关，而与 $x\in K$ 无关。

证明 先证明下面的结论：在定理的条件(i)下，则存在常数 $M>0$，对一切 $y\in R^n$，有

$$|Df(x)(y)|\leqslant M|y|,\quad \forall x\in K.$$

即常数 M 与 x 无关，对紧集 K 中的一切 x 有公共的 M。

设 $f=(f_1,f_2,\cdots,f_n)$，$y=(y_1,\cdots,y_n)$，那么

$$Df(x)(y)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x)\\ \cdots & \cdots & \cdots\\ \frac{\partial f_n}{\partial x_1}(x) & \cdots & \frac{\partial f_n}{\partial x_n}(x)\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

$$=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x)y_1+\cdots+\frac{\partial f_1}{\partial x_n}(x)y_n\\ \cdots\cdots\\ \frac{\partial f_n}{\partial x_1}(x)y_1+\cdots+\frac{\partial f_n}{\partial x_n}(x)y_n\end{pmatrix}.$$

所以

$$|Df(x)(y)|\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\frac{\partial f_i}{\partial x_j}(x)\right||y_i|,$$

再由 $\frac{\partial f_i}{\partial x_j}$ 在紧集 K 的连续性得知，存在常数 $c>0$，使

$$\left|\frac{\partial f_i}{\partial x_j}(x)\right|\leqslant c\quad (i,j=1,2,\cdots,n;x\in K).$$

于是

$$\begin{aligned}|Df(x)(y)|&\leqslant nc(|y_1|+\cdots+|y_n|)\\&\leqslant n^2c|y|\quad (y\in R^n;x\in K).\end{aligned}$$

继续证明下去：在定理的条件下，由逆映射定理知道，在每一个 $x\in D$ 的局部范围内，f^{-1} 是连续可微的，并且 $Df^{-1}(y)=(Df(x))^{-1}$，其中 $y=f(x)$。为了记号的简单起见，记 $Df(x)=l_x$，则 $Df^{-1}(y)=l_x^{-1}$。

在 f 的作用下，将 S 映射为 $f(S)$，因为 f^{-1} 是连续的，S 是开

集，所以 $f(S)$ 也是一个开集。设 $f(S)$ 在 l_x^{-1} 的作用下映射为 $l_x^{-1}(f(S))$，它仍旧是一个开集，于是

$$v(l_x^{-1}(f(S))) = |\det l_x^{-1}| v(f(S))$$

$$= \frac{1}{|\det l_x|} v(f(S)),$$

问题化为只要证明

$$\frac{v(l_x^{-1}f(S))}{v(S)} \to 1 \quad (h \to 0).$$

并且这一极限在 $x \in K$ 上是一致的。

先证明对任意给定的 ε $(0<\varepsilon<1)$，存在适当小的 $h>0$ 和两个 n 维矩形 A_1 和 A_2，A_1 的每个边的边长都是 $h(1-\varepsilon)$；A_2 的每个边的边长都是 $h(1+\varepsilon)$，使得

$$A_1 \subset l_x^{-1}(f(S)) \subset A_2.$$

证明如下：因为 f 在 D 连续可微，所以

$$f(x') - f(x) = Df(x)(x'-x) + r(x', x),$$

$$\lim_{|x'-x|\to 0} \frac{|r(x',x)|}{|x'-x|} = 0 \quad (\text{在紧集 } K \subset D \text{ 是一致的}).$$

将 l_x^{-1} 作用在等式 $f(x') - f(x) = Df(x)(x'-x) + r(x', x)$ 的两端，得

$$l_x^{-1}(f(x')) - l_x^{-1}(f(x)) = x' - x + l_x^{-1}(r(x', x)).$$

K 是紧集，从而 $f(K)$ 也是紧集（1.3 定理 3），当 $x \in K$ 时，$y = f(x) \in f(K)$，由刚才证明的结论，可以断言存在常数 $M>0$，对一切 $x \in K$（即 $y = f(x) \in f(K)$），有

$$|l_x^{-1}(r(x', x))| = |Df^{-1}(y)(r(x', x))| \leqslant M|r(x', x)|,$$

于是

$$\frac{|l_x^{-1}(r(x', x))|}{|x'-x|} \leqslant M\frac{|r(x',x)|}{|x'-x|} \to 0 \quad (|x'-x| \to 0),$$

这一极限关于 $x \in K$ 是一致的，也就是存在与 x 无关的 $h>0$，当 $|x'-x|<h$ 时，

$$|l_x^{-1}(f(x')) - l_x^{-1}(f(x)) - (x'-x)| < \varepsilon|x'-x|,$$

即

$$(1-\varepsilon)|x'-x|\leqslant|l_x^{-1}(f(x'))-l_x^{-1}(f(x))|\leqslant(1+\varepsilon)|x'-x|.$$

这样便证明了

$$A_1\subset l_x^{-1}(f(S))\subset A_2,$$

其中 A_1 和 A_2 就是前面所说的两个 n 维矩形。

再注意到

$$\frac{v(A_1)}{v(S)}\leqslant\frac{v(l_x^{-1}(f(S)))}{v(S)}\leqslant\frac{v(A_2)}{v(S)}$$

以及 $v(S)=h^n$, $v(A_1)=(1-\varepsilon)^nh^n$, $v(A_2)=(1+\varepsilon)^nh^n$, 所以当 $h\to0$ 时,

$$\frac{v(l_x^{-1}(f(S)))}{v(S)}\to1,$$

并且这一极限关于 $x\in K$ 是一致的。证毕。

习　　题

1. 设 $xy+yu+uv+vx=0$, $x^2+y^2+u^2+v^2=4$, 说明在 $(x, y, u, v)=(\sqrt{2}, 0, \sqrt{2}, 0)$ 的某个邻域内可以由方程组确定 u、v 是 x、y 的可微函数, 并求 $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$.

2. 证明一般情形下的隐函数定理(定理 5)。

3. 设 D 是 R^n 中的一个开集, $f:D\to R^m$, f 在 D 连续可微, 又设 $m\leqslant n$, $a\in D$, 并且 $f(a)=0$, 矩阵 $Df(a)$ 的秩是 m, 在这些条件下, 可以得出怎样的结论?为什么?

4. 设 $f:R^n\to R^n$ 并且连续可微, A 是 R^n 中的一个开集, 对每一个 $x\in A$, $\det Df(x)\neq0$, 又设 $y\notin f(A)$, 作

$$\psi(x)=|y-f(x)|^2,$$

证明 $D\psi(x)\neq0$, $x\in A$.

5. 设 $f:R^n\to R^n$ 并且连续可微, 对 R^n 中任何 x、y, 有

$$|f(y)-f(x)|\geqslant c|x-y|,$$

$c>0$ 是一个与 x、y 无关的常数。试证明

(i) f 是 R^n 上的单射;

(ii) $\det Df(x)\neq0$, $\quad x\in R^n$;

(iii) $f(R^n)=R^n$.

第3章　欧几里得空间上的积分

这一章的大多数内容都是读者在数学分析中已经熟知的，但在数学分析中却没有涉及到一个核心问题，那就是黎曼(Riemann)可积函数的特征是什么。我们常常说黎曼可积的函数是"不连续点不太多"的函数，但何谓"不太多"呢？本章将对这一问题作出明确的回答。同时，还将概括地叙述有关内容，并略去其中读者已经熟悉的那些证明。须要详细了解这些内容的读者可以参阅任何一本数学分析的教科书，例如复旦大学编的《数学分析》(欧阳光中，朱学炎，秦曾复编，上海科学技术出版社1982年出版)。

§3.1　可积函数的特征

可积

在n维欧几里得空间R^n上，我们先引进n维闭矩形上的积分，然后再引进一般的有界闭集上的积分，但后者须要对有界闭集加以限制。

在R^n内，称

$$
\begin{aligned}
A &= [a_1, b_1] \times \cdots \times [a_n, b_n] \\
&= \{(x_1, \cdots, x_n) \in R^n \mid a_i \leqslant x_i \leqslant b_i, i = 1, 2 \cdots, n\}
\end{aligned}
$$

是一个n维闭矩形(这一定义和2.3中的定义是相同的)，当$b_i - a_i$ $(i=1, 2, \cdots, n)$都相等时，称A是n维闭正方形，其容积$v(A)$和直径$\rho(A)$分别是

$$
\begin{aligned}
v(A) &= (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n), \\
\rho(A) &= \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}.
\end{aligned}
$$

设$p_i = \{t_0^i, t_1^i \cdots, t_{n_i}^i\}$是区间$[a_i, b_i]$的一个划分，即$a_i = t_0^i < t_1^i < \cdots < t_{n_i}^i = b_i$，则$p = \{p_1, \cdots, p_n\}$就是$A$的一个划分，它将$A$

划分为$n_1\times n_2\times\cdots\times n_n$个子矩形$\Delta_{i_1\cdots i_n}$，记$i=(i_1,\cdots i_n)$，这样就可以把$\Delta_{i_1i_2\cdots i_n}$简单的记为$\Delta_i$，其容积是$v(\Delta_i)$。再设$f$是定义在$A$上的有界实值函数，令

$$M_i=\sup\{f(x)\,|\,x\in\Delta_i\},$$
$$m_i=\inf\{f(x)\,|\,x\in\Delta_i\},$$

作f在A上对应于划分P的上和与下和

$$U(f,p)=\sum_i M_iv(\Delta_i),$$

$$L(f,p)=\sum_i m_iv(\Delta_i),$$

它们具有以下性质：

(i) 对任何划分，$L(f、P)\leqslant U(f,P)$。

(ii) 设划分P'是划分P的加细，则
$U(f,P)\geqslant U(f,P')$，$L(f、P)\leqslant L(f,P')$。

(iii) 对任意两个划分P和Q，都有

$$L(f,P)\leqslant U(f,Q).$$

(iv) 对给定的有界函数，数集$\{U(f,P)\}$和$\{L(f,P)\}$都是有界数集。

因此，

$$\overline{\int}_A f=\inf\{U(f,P)\,|\,\text{一切划分 }P\},$$

$$\int_{-A} f=\sup\{L(f,P)\,|\,\text{一切划分 }P\},$$

对任何给定的有界函数f总是存在的，分别称它们是f在A上的上积分和下积分。但两者可能不相等。如果两者相等，就称f在A是黎曼可积的，简称它是可积。并将$\overline{\int}_A f$或$\int_{-A} f$记为$\int_A f$，称这个数是f在A上的积分，有时写为$\int_A\cdots\int f(x_1,\cdots,x_n)dx_1\cdots dx_n$。

定理1 设f是定义在n维闭矩形A上的一个有界函数，则f在A可积的充分和必要条件是：对任意给定的$\varepsilon>0$，存在A的一个划分P，使得

$$U(f,P)-L(f,P)<\varepsilon.$$

利用这一条件立即可以证明闭矩形上的连续函数一定是可积的。

根据“积分是黎曼和的极限”这一事实，设 $P=\{p_1,p_2,\cdots,p_n\}$ 是 A 的一个划分，将 A 划分为一列子矩形 Δ_i $(i=1,2,\cdots,m)$。记 $\|P\|$ 是所有子矩形的最大直径，即 $\|P\|=\max\{\rho(\Delta_i)\,|\,i=1,2,\cdots,m\}$，称 $\|P\|$ 是划分 P 的范数，作

$$\sigma=(f,p,\xi)=\sum_{i=1}^{m}f(\xi_i)v(\Delta_i),$$

其中 $\xi_i\in\Delta_i$，称 $\sigma(f,p,\xi)$ 是 f 在 A 上对应于划分 P 和 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)$ 的黎曼和，它不仅与划分 P 有关，而且还与 ξ 有关。如果

$$\lim_{\|P\|\to 0}\sigma(f,p,\xi)=\lim_{\|P\|\to 0}\sum_{i=1}^{m}f(\xi_i)v(\Delta_i)$$

存在，并且这一极限与划分 P 以及 ξ_i 在 Δ_i 内的选取无关，就称 f 在 A 黎曼可积，并称这一极限是 f 在 A 的积分，记为 $\int_A f$。这个定义和我们前面所说的定义是否一致呢？

定理 2 下列五个命题是彼此等价的：

(i) $\overline{\int}_A f=\underline{\int}_A f$.

(ii) 对任意给定的 $\varepsilon>0$，存在 A 的一个划分 P，

$$U(f,P)-L(f,P)<\varepsilon.$$

(iii) 对任意给定的 $\varepsilon>0$，存在 $\delta>0$，对一切 $\|P\|<\delta$ 的划分 P，有

$$U(f,P)-L(f,P)<\varepsilon.$$

或者写为

$$\lim_{\|P\|\to 0}(U(f,P)-L(f,P))=0.$$

(iv) $$\lim_{\|P\|\to 0}U(f,P)=\lim_{\|P\|\to 0}L(f,P).$$

(v) $$\lim_{\|P\|\to 0}\sigma(f,P,\xi)=\lim_{\|P\|\to 0}\sum_{i=1}^{m}f(\xi_i)v(\Delta_i)$$

存在，并且这一极限与划分P和ξ无关。

此外，(iv)和(v)中的极限值是相同的，并且都等于(i)中的积分。

这些等价命题从表面的形式来看，(ii)最弱，而(iv)或(v)最强(事实上它们是一样的强弱)，因此当我们不知道f是否可积而须要判断它的可积性时，利用(ii)要方便些，但当我们已知f是可积的，须要进一步研究它时，(iv)或(v)将给我们提供更多的信息。

引进了积分的概念和给出了可积的充要条件之后，自然会问怎样的函数才可积呢?回答是：闭矩形上的连续函数必可积，但“连续”的条件还可以减弱，我们将在下面深入研究这个问题。

零容度集

设S是R^n内的一个子集，如果对任意给定的$\varepsilon>0$，存在有限个n维闭矩形Δ_1，Δ_2，…，Δ_k满足(i) Δ_1，Δ_2，…，Δ_k覆盖了S，(ii)$\Delta_i(i=1,2,\cdots,k)$的容积之和小于ε。亦即

(i) $$S\subset\bigcup_{i=1}^{k}\Delta_i;$$

(ii) $$\sum_{i=1}^{k}v(\Delta_i)<\varepsilon.$$

则称S是一个n维零容度集，或者称S的容度是0。例如直线上有限个点组成的集显然是一个1维零容度集，由点$\frac{1}{n}$ $(n=1, 2, \cdots)$组成的集也是1维零容度集，但在区间$[0, 1]$内的所有有理数组成的集不是零容度集。在R^n中，一个n维闭矩形的边界是n维零容度集，这是因为n维闭矩形的边界由$2n$个$n-1$维的闭矩形组成，而每一个$n-1$维闭矩形总可以用一个容积充分小的n维闭矩形来包含它。

我们知道，一个定义在n维闭矩形A上的有界函数，如果它的所有不连续点组成的集是一个零容度集，那么它必在A可积，然而反过来，一个在A可积的函数，它在A内的所有不连续点可能不是

一个零容度集，而比零容度集更大一些，例如在区间 $[0,1]$ 上的黎曼函数

$$f(x)=\begin{cases}\frac{1}{p}, \text{当 } x \text{ 是正有理数 } x=\frac{q}{p}\text{(既约分数)}; \\ 0, \text{当 } x \text{ 是无理数或 } 0.\end{cases}$$

它在 $[0,1]$ 可积，但 f 在 $[0,1]$ 内的不连续点是所有非 0 的有理数组成的集，它不是零容度集。因此，“不连续点是零容度集的函数”还不足以刻划黎曼可积的特征。有没有关于黎曼可积的完善的结论呢？有的！这将牵涉到零测度集的概念。

零测度集

引进零测度集的概念必须利用“可列”这一数学术语。设 A 是一个集，N 是所有自然数组成的集：

$$N=\{1,2,\cdots,n,\cdots\},$$

如果 A 的元素和 N 的元素之间可以建立一一对应关系，也就是说如果存在映射 $\varphi:N\to A$，并且 φ 是双射，则称 A 是一个可列集。又称 A 中的元素有可列无限多个。直观上看，可列集的元素总可以按标号 1, 2, …, n, …排列出来：

$$A=\{a_1, a_2, \cdots, a_n, \cdots\},$$

其中 a_n 和自然数 n 对应。由一一对应的特点知道，可列集的元素和自然数“一样多”，是可列无限多。

例 1 所有整数组成的集 Z 是可列集。

证明 由

$$Z=\{0, 1, -1, 2, -2, \cdots, n, -n, \cdots\}$$

建立如下的一一对应关系：

$$\begin{array}{ccccccccc} 0 & 1 & -1 & 2 & -2 & \cdots & n & -n\cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \updownarrow \\ 1 & 2 & 3 & 4 & 5 & \cdots & 2n & 2n+1\cdots \end{array}$$

或者说 Z 中的元素可以用标号排列出来，例如元素 n 是 $2n$ 号，

$-n$ 是 $2n+1$ 号($n=1, 2, \cdots$)，(当然不是按大小顺序排列)，可见 Z 是一个可列集，证毕。

这个例子表明：虽然 N 是 Z 的一个真子集，但两者的元素却"一样多"，这正是无限集和有限集的根本差别。

例 2 设 A、B 都是可列集，则 $A\cup B$ 也是可列集。

证明 设

$$A=\{a_1, a_2, \cdots, a_n, \cdots\};$$
$$B=\{b_1, b_2, \cdots, b_n, \cdots\}.$$

则可以把 $A\cup B$ 中的元素排列为

$$A\cup B=\{a_1, b_1, a_2, b_2, \cdots, a_n, b_n, \cdots\}$$

(在 $A\cup B$ 中要去掉重复出现的元素)，可见 $A\cup B$ 是可列集。证毕。

推论 有限个可列集的并集是可列集。

例 3 所有有理数组成的集 Q 是可列集（这是一个惊人的事实，它说明有理数和自然数"一样多"）。

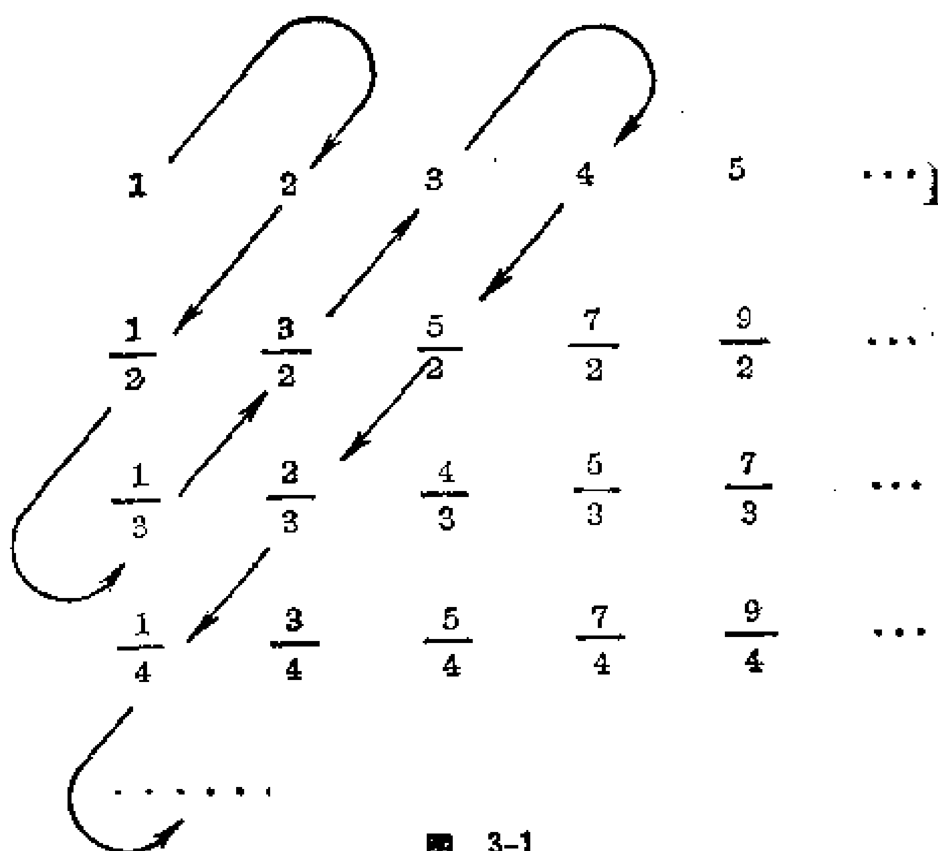


图 3-1

证明 设Q^+是所有正有理数组成的集；Q^-是所有负有理数组成的集。先证明Q^+是可列集。将Q^+的元素列成上面的方阵(见图3-1)，第一行是分母为1的所有正有理数(即所有自然数)，按大小顺序排列，第二行是分母为2的所有正的既约分数，也按大小顺序排列，第三行是分母为3的所有正的既约分数，按大小顺序排列，…，等等，如图3-1所示。然后按箭头所表示的次序将所有正有理数排列为

$$Q^+=\left\{1,\ 2,\ \frac{1}{2},\ \frac{1}{3},\ \frac{3}{2},\ 3,\ 4,\ \frac{5}{2},\cdots\right\},$$

所有正有理数无一挂漏，所以Q^+是一个可列集。

同样，Q^-也是一个可列集。再由

$$Q=Q^+\cup Q^-\cup\{0\}$$

和例2得Q是可列集。证毕。

例4 设$A_1,\ A_2,\ \cdots,\ A_n,\ \cdots$都是可列集，那么$\bigcup\limits_{n=1}^{\infty}A_n$也是可列集，即可列无限多个可列集的并集仍旧是可列集。

证明 设

$$A_1=\{a_{11},\ a_{12},\ a_{13},\ \cdots\},$$
$$A_2=\{a_{21},\ a_{22},\ a_{23},\ \cdots\},$$
$$A_3=\{a_{31},\ a_{32},\ a_{33},\ \cdots\},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

然后再按照例3的方法将$\bigcup\limits_{n=1}^{\infty}A_n$中的元素排列出来，证毕。

是不是任何无限集一定都是可列的呢？

例5 $[0,1]=\{x\in R\,|\,0\leqslant x\leqslant 1\}$是不可列集。

证明 采用反证法，假设$[0,\ 1]$是可列集，则

$$[0,\ 1]=\{x_1,\ x_2,\ \cdots,\ x_n,\ \cdots\}.$$

将区间$[0,\ 1]$三等分，其中必有一个子区间不含x_1，记这个子区间是$[a_1,\ b_1]$，$x_1\notin[a_1,\ b_1]$。将$[a_1,\ b_1]$三等分，其中必有一个子区间不含x_2，记它是$[a_2,\ b_2]$，$x_2\notin[a_2,\ b_2]$。将$[a_2,\ b_2]$三等分，…，得一列闭区间$[a_n,\ b_n]\ (n=1,\ 2,\ \cdots)$满足：

(i) $[a_1, b_1]\supset[a_2, b_2]\supset\cdots\supset[a_n, b_n]\supset\cdots$

$$b_n-a_n\to 0\ (n\to\infty),$$

由区间套定理，存在$\xi\in[a_n, b_n]\,(n=1, 2, \cdots)$，显然$\xi\in[0, 1]$。

(ii) $x_n\notin[a_n, b_n]\,(n=1, 2, \cdots)$。

由(i)和(ii)得$\xi\neq x_n\ (n=1, 2, \cdots)$，但$\xi\in[0,1]$与$[0, 1]=\{x_1, x_2, \cdots, x_n, \cdots\}$矛盾。证毕。

现在引进零测度集的概念。设$S\subset R^n$，如果对任意给定的$\varepsilon>0$，存在可列无限多个n维闭矩形$\Delta_1, \Delta_2, \cdots, \Delta_n, \cdots$，使

$$S\subset\bigcup_{n=1}^{\infty}\Delta_n,\quad \sum_{n=1}^{\infty}v(\Delta_n)<\varepsilon,$$

就称S是R^n中的一个零测度集。显然零容度集一定是零测度集，但反之不然。容易知道，在零测度集的定义中，$S\subset\bigcup\limits_{n=1}^{\infty}\Delta_n$可以换为$S\subset\bigcup\limits_{n=1}^{\infty}(\Delta_n$的内部)，也就是说，可以把闭矩形换为开矩形。

例6 直线上所有正整数组成的集N是零测度集，

证明 对任意给定的$\varepsilon>0$，取闭区间$\Delta_n=\left[n-\frac{1}{4}\cdot\frac{\varepsilon}{2^n},\ n+\frac{1}{4}\cdot\frac{\varepsilon}{2^n}\right]\,(n=1, 2, \cdots)$，那么$\{\Delta_n\}$覆盖了$N$，并且它们的长度之和

$$\sum_{n=1}^{\infty}v(\Delta_n)=\frac{\varepsilon}{2}\sum_{n=1}^{\infty}\frac{1}{2^n}<\varepsilon.$$

证毕

很明显，N不是零容度集。

例7 直线上的所有有理数组成的集Q是直线上的零测度集，平面上所有有理点组成的集$Q^2=\{(x, y)\in R^2\,|\,x, y$都是有理数$\}$是平面上的零测度集。更一般的，$R^n$内的任何可列集都是$n$维零测度集。

证明如同例6那样，因为Q和Q^2都是可列集，将其元素排列出来，再按照例6的方法即得出证明。例如设

$$Q=\{r_1, r_2, \cdots, r_n, \cdots\},$$

对任意给定的$\varepsilon>0$，取一列闭区间是$\left[r_n-\frac{1}{4}\cdot\frac{\varepsilon}{2^n},\ r_n+\frac{1}{4}\cdot\frac{\varepsilon}{2^n}\right]$

$(n=1, 2, \cdots)$，就可以得出 Q 是零测度集。

例 8 可列无限多个零容度集的并集是零测度集。

证明是明显的.

更一般的还可以证明可列无限多个零测度集的并集是零测度集(习题 4)。此外，明显地有：零测度集的任何子集都是零测度集。

振幅

设 E 是 R^n 中的一个子集，f 是定义在 E 上的有界实值函数，记

$$\omega(f, E)=\sup\{f(x)\,|\,x\in E\}-\inf\{f(x)\,|\,x\in E\},$$

称 $\omega(f, E)$ 是 f 在 E 上的振幅。容易知道如果 $E'\subset E$，那么

$$\omega(f, E')\leqslant\omega(f, E).$$

现在，我们利用 f 在 E 的振幅来定义 f 在点 x 的振幅。对任意给定的 $\varepsilon>0$，设

$$A_\varepsilon=[x_1-\varepsilon, x_1+\varepsilon]\times\cdots\times[x_n-\varepsilon, x_n+\varepsilon],$$

它是以点 x 为中心的 n 维闭矩形。令 ε 单调趋于 0，定义

$$\omega(f, x)=\lim_{\varepsilon\to 0}\omega(f, A_\varepsilon\cap E),$$

由于 $\omega(f, A_\varepsilon\cap E)$ 关于 ε 是单调的，所以这一极限存在，称它是 f 在点 x 的振幅

例 9 设 f，g 都是一元实值函数，

$$f(x)=\begin{cases}1, & x\leqslant 0;\\ -x, & x>0.\end{cases}$$

$$g(x)=\begin{cases}\sin\dfrac{1}{x}, & x\neq 0;\\ 0, & x=0.\end{cases}$$

容易验证

$$\omega(f, 0)=1,\quad \omega(g, 0)=2.$$

由定义立即知道 $\omega(f, x)\geqslant 0$。又如果 x 是闭矩形 A 的内点，那么 $\omega(f, A)\geqslant\omega(f, x)$。

定理 3 f 在点 x 连续当且仅当 $\omega(f, x)=0$。

证明是明显的。

定理 4 设 $f: R^n \to R$，对任何给定的 $c>0$，设

$$E_c = \{x \in R^n \mid \omega(f, x) \geqslant c\},$$

则 E_c 是闭集。

证明 设 x 是 E_c 的一个聚点(极限点)，要证明的是：$x \in E_c$ 即 $\omega(f, x) \geqslant c$。由于 x 是 E_c 的聚点，所以存在一列 $x_n \in E_c (n = 1, 2, \cdots)$，$x_n \to x$。对任何一个以点 x 为中心的闭矩形 A，在 $\{x_n\}$ 中必有一个 x_n 是 A 的内点，所以

$$\omega(f, A) \geqslant \omega(f, x_n) \geqslant c,$$

由 A 的任意性，得 $\omega(f, x) \geqslant c$，证毕。

定理 5 设 f 是定义在 n 维闭矩形 A 上的有界函数，并且对每一个点 $x \in A$，$\omega(f, x) < c$，则存在 A 的一个划分 P，使

$$U(f, P) - L(f, P) < c \cdot v(A).$$

证明 对每一个 $x \in A$，$\omega(f, x) < c$，由振幅的定义可知道，存在一个以 x 为中心的 n 维闭矩形 I_x 满足

$$\omega(f, I_x \cap A) < c.$$

将闭矩形 I_x 的边界去掉，就成为一个含 x 的开矩形，而每一个 $x \in A$ 都有这样一个开矩形，根据海涅-波莱尔 (Heine-Borel) 定理，存在有限个这样的闭矩形覆盖 A，记它们是 $I_1, I_2, \cdots, I_m$。

作 A 的一个划分 P，使得在划分 P 下的每一个子矩形都含在某一个 $I_i (i = 1, 2, \cdots, m)$ 内，则

$$U(f, P) - L(f, P) < c \cdot v(A)。证毕。$$

可积函数的特征

定理 6 设 f 是定义在 n 维闭矩形 A 上的有界函数。则 f 在 A 可积(黎曼意义下)的充分和必要条件是：f 在 A 的所有不连续点组成的集是 n 维零测度集。

定义在 A 上的函数 f，如果它的所有不连续点组成一个零测度集，就称 f 在 A 上几乎处处连续，这时定理 6 又可以表达为：设 f 在 A 有界，则 f 在 A 黎曼可积的充分和必要条件是：f 在 A 上几

乎处处连续。

证明 先证明必要性：设 f 在 A 可积，令

$$E=\{x\in A\mid f\text{ 在 }x\text{ 不连续}\},$$

我们的目的是证明 E 是一个 n 维零测度集。注意到下列两个命题是等价的：

(i) $x\in E$，即 x 是 f 的一个不连续点；

(ii) $\omega(f, x)>0$。

所以只要证明 $E=\{x\in A\mid \omega(f, x)>0\}$ 是一个零测度集。而

$$E=\bigcup_{m=1}^{\infty}E_m,\qquad E_m=\left\{x\in A\mid \omega(f, x)\geqslant\frac{1}{m}\right\},$$

再由例 8，只要证明每一个 E_m 是零容度集就可以了。

固定 m，由 f 在 A 可积知道：对任意给定的 $\varepsilon>0$，存在 A 的一个划分 P，它将 A 划分为 k 个子矩形 I_1, …, I_k，使

$$U(f, P)-L(f, P)<\frac{\varepsilon}{m}.$$

E_m 中的点 x 可以分为两类：(1) x 正好在某个子矩形的边界上，记这种点的全体是 E_m^1，由于在划分 P 之下的子矩形是有限个，而每一个 n 维矩形的边界是一个 n 维零容度集，所以 E_m^1 是一个零容度集。(2) 其余的点，其全体记为 E_m^2，即 $E_m^2=E_m-E_m^1$，E_m^2 中任何一点 x 必定是某个子矩形 I_i 的内点，因此

$$\omega(f, I_i)\geqslant\omega(f, x)\geqslant\frac{1}{m},$$

令 I_{i_1}, …, I_{i_s} 是含有 E_m^2 中的点的所有子矩形，则

$$E_m^2\subset I_{i_1}\cup\cdots\cup I_{i_s},$$

并且又有

$$\begin{aligned}&\frac{1}{m}(v(I_{i_1})+\cdots+v(I_{i_s}))\\&\leqslant\omega(f, I_{i_1})v(I_{i_1})+\cdots+\omega(f, I_{i_s})v(I_{i_s})\\&\leqslant\sum_{i=1}^{k}\omega(f, I_i)v(I_i).\\&=U(f, P)-L(f, P)<\frac{\varepsilon}{m},\end{aligned}$$

这样便得到

$$v(I_{i_1})+\cdots+v(I_{i_r})<\varepsilon,$$

即 E_m^2 是一个零容度集，于是 $E_m=E_m^1\cup E_m^2$ 是一个零容度集，必要性证毕。

再证明充分性：设

$$E=\{x\in A\mid\omega(f,x)>0\}$$

是一个零测度集。

第一步：对任意给定的 $\varepsilon>0$，存在正整数 m，使 $\dfrac{v(A)}{m}<\varepsilon$。再设 E_m 同前，由假设知道 E_m 是一个零测度集，因此存在一列 n 维闭矩形 $I_1, I_2, \cdots, I_n, \cdots$，使

$$E_m\subset\bigcup_{n=1}^{\infty}(I_n\text{ 的内部}),$$

$$\sum_{n=1}^{\infty}v(I_n)<\frac{\varepsilon}{\omega(f,A)}.$$

E_m 是闭集（定理 4），同时又是有界集，由海涅-波莱尔定理，在 $I_1, I_2, \cdots$ 中存在有限个 $I_{n_1}, I_{n_2}, \cdots, I_{n_k}$ 覆盖 E_m。

第二步：在 A 中将这些 $I_{n_j}(j=1, 2, \cdots, k)$ 去掉，剩下来的部分的闭包总可以分解为有限个闭矩形 $J_1, J_2, \cdots, J_l$ 的并，并且这些闭矩形的内部两两不相交，于是

$$A=J_1\cup\cdots\cup J_l\cup I_{i_1}\cup\cdots\cup I_{i_k}.$$

每一个 $J_j\ (j=1, 2, \cdots, l)$ 的内部都不含有 E_m 中的点，由定理 5，存在 J_j 的一个划分 P_j，使在 J_j 上有

$$U(f,P_j)-L(f,P_j)<\frac{1}{m}v(J_j).$$

在划分 P_j 之下，J_j 被划分为有限多个闭子矩形，不同的 j 得出另外有限多个闭子矩形，记这些闭子矩形的全体是 D。

第三步：作 A 的一个划分 P，要求它满足下面的条件：在划分 P 下，子矩形可分为两类：一类是含在 D 中的某个子矩形内；另一类是含在某个 I_{n_j} 内，于是，在这个划分之下就有

$$U(f,P)-L(f,P)\leqslant\sum_{j=1}^{l}(U(f,P_j)-L(f,P_j))$$

$$+\sum_{i=1}^{k}\omega(f, I_{n_i})v(I_{n_i})$$

$$<\frac{1}{m}\sum_{j=1}^{l} v(J_j) + \omega(f, A) \sum_{i=1}^{k} v(I_{n_i})$$

$$<\frac{1}{m} v(A) + \omega(f, A)\frac{\varepsilon}{\omega(f, A)}<2\varepsilon.$$

这就证明了 f 在 A 可积，证毕。

有界集上的积分

如何把闭矩形上积分的概念拓广到任意一个有界集上去呢？设 D 是 R^n 中的一个有界集，f 是定义在 D 上的有界函数。令 A 是包含 D 的一个闭矩形，作

$$g(x)=\begin{cases} f(x), & x\in D; \\ 0, & x\in A-D. \end{cases}$$

如果 g 在 A 可积，就称 f 在 D 可积，并定义 f 在 D 的积分为

$$\int_D f=\int_A g.$$

从形式上看，这一定义似乎和闭矩形 A 的选取有关，但事实上却不然，留给读者考虑(作为习题)。

接下来的问题自然是在什么条件下 f 在 D 可积？能不能回答说：只要 f 在有界集 D 连续(或者几乎处处连续)，就保证 f 在 D 可积呢？考察一个最好的情况：设在 D 内 $f(x)=c$，c 是常数，$c\neq 0$。A 是包含 D 的闭矩形，g 的定义如上面所说，这时 D 的边界点必定是 g 的不连续点，并且 g 在 A 上所有不连续的点组成的集合就是 D 的边界，定理 6 表明如果 D 的边界不是零测度集，则 g 在 A 不可积，从而 f 在 D 不可积，因此 f 在 D 的可积性不仅依赖于函数 f 本身的性质，还依赖于 D 的边界的性质。确实存在这样的有界集，它的边界不是零测度集，例如在 $[0, 1]$ 内所有有理数组成的集，它的边界是 $[0, 1]$，显然不是零测度集。甚至一个有界开集，它的边界也可能不是零测度集(习题 8)。

由定理 6 立即得：

定理 7 设D是R^n内的一个有界集，D的边界是n维零测度集，f是定义在D内的有界的几乎处处连续的函数，那么f在D可积。

在黎曼积分的应用中，D的边界性质常常是很好的。以平面情形为例，D的边界大多是由有限条光滑曲线段或者由有限条可求长的曲线段组成，这样的边界不仅是零测度集，而且还是零容度集。

习　题

1．(i) 设$P_n(x)=a_0+a_1x+\cdots+a_nx^n$是$n$次整系数多项式，即系数$a_i(i=0,1,\cdots,n)$都是整数，证明这种多项式的全体组合的集是一个可列集(其中n是给定的一个正整数)。

(ii) 证明所有整系数多项式组成的集是一个可列集。

(iii) 实数a如果是某一个整系数多项式的根，就称它是一个代数数，证明所有代数数组成的集是一个可列集。

2．证明$Q^2=\{(x,y)\in R^2|x,y$都是有理数$\}$是R^2内的零测度集，但不是零容度集。

3．设S是R^n中的可列集，只有有限个聚点(极限点)，证明S是零容度集。

4．证明可列无限多个零测度集的并集是零测度集。

5．设f在闭区间$[a,b]$上可积，证明f的连续点在$[a,b]$内稠密(利用定理6怎样证明？不利用定理6怎样证明？)。

6．在有界集D上的积分定义中，证明f在D的可积性及其积分与闭矩形A的选取无关。

7．设S是R^n内的有界集，定义S上的特征函数χ_s如下：

$$\chi_s(x)=\begin{cases}1, & x\in S;\\ 0, & x\notin S.\end{cases}$$

设S是R^n中的零容度集，证明χ_s在S可积，并且$\int_S\chi_s=0$. 如果S是R^n的零测度集，问χ_s在A是否可积？

8．设Q是区间$(0,1)$内所有有理数组成的集。

(i) 证明对任意给定的$e<0$，$e<1$，存在$(0,1)$内的一列开区间(α_n,β_n)

$(n=1,2,\cdots)$ 覆盖 Q，并且 $\sum_{n=1}^{\infty}(\beta_n-\alpha_n)<\varepsilon$。

(ii) 令 $S=\bigcup_{n=1}^{\infty}(\alpha_n,\beta_n)$、证明 S 的边界是 $[0,1]-S$。

(iii) S 是开集，它的边界是零测度集吗？

§3.2 积分的性质

基本性质

积分具有一些重要的性质，这些性质是读者早已熟知的，列出如下，在这里利用定理6来证明某些结论，这样做比在数学分析中的证明简单得多。

(1) 线性 设函数 f，g 都在闭矩形 A 可积，α 和 β 是任意两个实数，则 $\alpha f+\beta g$ 也在 A 可积，并且

$$\int_A(\alpha f+\beta g)=\alpha\int_A f+\beta\int_A g.$$

现在利用定理6来证明 $\alpha f+\beta g$ 在 A 可积。设 f，g 都在点 x 连续，则 $\alpha f+\beta g$ 也在点 x 连续，这表明若设

$$E=\{x\in A\,|\,\alpha f+\beta g\text{ 在 }x\text{ 不连续}\};$$

$$E_1=\{x\in A\,|\,f\text{ 在 }x\text{ 不连续}\};$$

$$E_2=\{x\in A\,|\,g\text{ 在 }x\text{ 不连续}\}.$$

则有

$$E\subset E_1\cup E_2.$$

再由 f 和 g 的可积性得 $E_1\cup E_2$ 是零测度集，所以 E 是零测度集，即 $\alpha f+\beta g$ 在 A 可积。

(2) 设 f，g 都在 A 可积，并且 $f(x)\leqslant g(x)$，$\forall x\in A$，则

$$\int_A f\leqslant\int_B g.$$

(3) 设 f 在 A 可积，则 $|f|$ 也在 A 可积，并且

$$\left|\int_A f\right|\leqslant\int_A|f|.$$

利用定理6很容易获得 $|f|$ 的可积性，这是因为 f 的连续点必

定是$|f|$的连续点，所以

$$\{x\in A \mid |f| 在 x 不连续\} \subset \{x\in A \mid f 在 x 不连续\},$$

这表明当f是几乎处处连续时，$|f|$也是几乎处处连续的。

(4) 设f, g都在A可积，则$f\cdot g$也在A可积。

利用定理6也是很容易证明的。留作习题请读者自证。

(5) 中值定理　设f, g都在A可积，g在A上不改变符号，则存在常数μ，$m\leqslant\mu\leqslant M$(其中m, M分别是f在A上的下确界和上确界)使

$$\int_A fg=\mu\int_A g.$$

下面讨论累次积分的问题：设A, B分别是R^n和R^m中的闭矩形，f是定义在$A\times B$上的一个有界函数，任意固定$x\in A$，定义函数$g_x: B\to R$是

$$g_x(y)=f(x, y),\quad y\in B.$$

显然g_x是B上的一个有界函数，记它的上积分和下积分分别是：

$$u(x)=\overline{\int_B} g_x=\overline{\int_B} f(x, y)dy;$$

$$l(x)=\int_{-B} g_x=\int_{-B} f(x, y)dy,$$

它们都存在，这表明u和l都是定义在A上的函数，并且$l(x)\leqslant u(x)$, $x\in A$.

(7) 累次积分定理　在上面给出的假设和记号下，如果f在$A\times B$可积，则u和l都在A可积，并且

$$\int_{A\times B} f=\int_A u=\int_A\left\{\overline{\int_B} f(x, y)dy\right\}dx,$$

$$\int_{A\times B} f=\int_A l=\int_A\left\{\int_{-B} f(x, y)dy\right\}dx,$$

特别是：如果g_x在B可积，即

$$\overline{\int_B} f(x, y)dy=\int_{-B} f(x, y)\,dy=\int_B f(x, y)dy,$$

则

$$\int_{A\times B} f=\int_A\left\{\int_B f(x,\ y)dy\right\}dx.$$

这就是化 $A\times B$ 上的积分为先对 y 后对 x 的累次积分。

证明　设 P_A 是 A 的一个划分，P_B 是 B 的一个划分，它们分别将 A 和 B 划分为下列 k 个和 l 个闭子矩形：

$$I_1,\ I_2,\ \cdots,\ I_k \text{ 和 } J_1,\ J_2,\ \cdots,\ J_l,$$

这样，便得到 $A\times B$ 的一个划分 P，在这一划分下将 $A\times B$ 划分为 $k\times l$ 个闭子矩形，它们是：

$$I_i\times J_j\quad (i=1,\ 2,\ \cdots,\ k;\ j=1,\ 2,\ \cdots,\ l).$$

记

$$M_{ij}=\sup\{f(x,\ y)\,|\,x\in I_i,\ y\in J_j\};$$
$$m_{ij}=\inf\{f(x,\ y)\,|\,x\in I_i,\ y\in J_j\};$$
$$M_j(x)=\sup\{f(x,\ y)\,|\,y\in J_j\};$$
$$m_j(x)=\inf\{f(x,\ y)\,|\,y\in J_j\},$$

再注意到 $v(I_i\times J_j)=v(I_i)v(J_i)$，则

$$U(f,\ P)=\sum_{i=1}^{k}\ \sum_{j=1}^{l} M_{ij}v(I_i\times J_j)$$
$$=\sum_{i=1}^{k}\left\{\sum_{j=1}^{l} M_{ij}v(J_j)\right\}v(I_i);$$
$$L(f,\ P)=\sum_{i=1}^{k}\ \sum_{j=1}^{l} m_{ij}\,v(I_j\times J_j)$$
$$=\sum_{i=1}^{k}\ \left\{\sum_{j=1}^{l} m_{ij}\,v\,(J_j)\ \right\}v(I_i),$$

当 $x\in I_i$ 时，由于 $I_i\subset I_i\times J_j$，所以

$$M_i(x)\leqslant M_{ij},\ m_i(x)\geqslant m_{ij},\ x\in I_i,$$

于是(当 $x\in I_i$ 时)

$$\sum_{j=1}^{l}M_{ij}v(J_j)\geqslant\sum_{j=1}^{l}M_j(x)v(J_j)$$
$$\geqslant\overline{\int}_B g_x=u(x);$$
$$\sum_{j=1}^{l}m_{ij}\,v\,(J_j)\leqslant\sum_{j=1}^{l}m_j(x)v(J_j)$$

$$\leqslant \int_{-B} g_x = l(x),$$

上面两个不等式对一切 $x \in I_i$ 成立，所以

$$\sum_{j=1}^{l} M_{ij}\, v(J_j) \geqslant \sup\{u(x) \mid x \in I_i\};$$

$$\sum_{j=1}^{l} m_{ij}\, v(J_j) \leqslant \inf\{l(x) \mid x \in I_i\};$$

$$U(f, P) = \sum_{i=1}^{k} \left\{ \sum_{j=1}^{l} M_{ij}\, v(J_j) \right\} v(I_i) \geqslant U(u, P_A);$$

$$L(f, P) = \sum_{i=1}^{k} \left\{ \sum_{j=1}^{l} m_{ij}\, v(J_j) \right\} v(I_i) \leqslant L(l, P_A).$$

于是

$$L(f, P) \leqslant L(l, P_A) \leqslant U(l, P_A) \leqslant U(u, P_A) \leqslant U(f, P);$$

$$L(f, P) \leqslant L(l, P_A) \leqslant L(u, P_A) \leqslant U(u, P_A) \leqslant U(f, P),$$

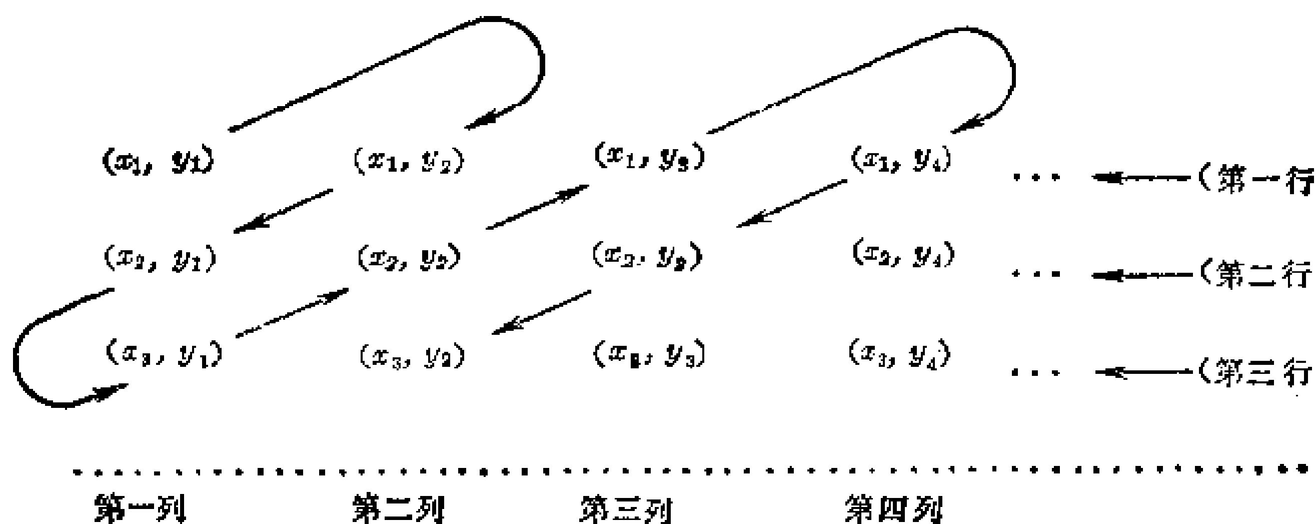
再由 f 的可积性得到 u 和 l 都在 A 可积，并且

$$\int_{A \times B} f = \int_A u = \int_A l.$$

证毕

读者已经知道这样一些例子：重积分存在但累次积分不存在，或者某个累次积分存在但重积分不存在。现在再举一个例子：二重积分不存在但两个累次积分存在并且相等，作出这个例子的关键在于构造出满足下列要求的平面点集。

设 $A = [0, 1] \times [0, 1]$，构造一个点集 $S \subset A$，满足：(i) S 在



A 内稠密；(ii) A 内每一条和 x 轴平行的直线如果和 S 相交，则交点只有一个；同样，A 内每一条和 y 轴平行的直线如果和 S 相交，交点也只有一个。

构造如下：A 中所有有理点组成的集 $\{(x, y)\in A \mid x, y$ 都是有理数$\}$ 是可列集，按下列方式（箭头）将其元素排列出来，如上一页的表所示。

将 A 四等分（图 3-2），按表中箭头顺序取出第一个属于 A_1 的点 (x_i, y_j)，然后将表中第 i 行和第 j 列的所有元素去掉，去掉的元素意味着“不存在了”，再从头开始按箭头顺序取出第一个属于 A_2 的点 (x_l, y_k)，然后将表中第 l 行和第 k 列的所有元素去掉，用同样的方式分别取出属于 A_3 和 A_4 中的点。

A_3	A_4
A_1	A_2

图 3-2

再将 $A_i (i=1, 2, 3, 4)$ 四等分，分为 $A_{i1}, A_{i2}, A_{i3}, A_{i4}$，按照刚才的方式分别取出属于 $A_{ij} (j=1, 2, 3, 4)$ 的点。再将 A_{ij} 四等分，…，如此继续下去，所有取出来的点组成的集记为 S，这就是所要求的集（为什么？证明是明显的）。

现在定义 A 上的函数 f 是

$$f(x, y)=\begin{cases}1, & \text{当 } (x, y)\in S; \\ 0, & \text{当 } (x, y)\notin S.\end{cases}$$

f 在 A 是不可积的，这是因为对 A 的任何划分 P，

$$U(f, P)=1, \quad L(f, P)=0.$$

但对任何固定的 $y\in[0, 1]$ 与任何固定的 $x\in[0, 1]$，分别有

$$\int_0^1 f(x, y)dx=0; \quad \int_0^1 f(x, y)dy=0,$$

所以两个累次积分

$$\int_0^1\left\{\int_0^1 f(x, y)dx\right\}dy=\int_0^1\left\{\int_0^1 f(x, y)dy\right\}dx=0.$$

性质(1)～(7)都是在闭矩形 A 上讨论的，但它们都可以很自然的拓广到任何具有零测度边界的有界集上，这里不一一论述了。

(8) 变量代换　设D是R^n内的具有零测度边界的有界集，f在D连续。又设变换$T:D'\to D$是连续可微的双射，并且

$$J=\det DT(t)\neq 0,\quad \forall t\in D'.$$

那么

$$\int_D f=\int_{D'} f\circ T\cdot|J|.$$

设D中的点是$(x_1,\ \cdots,\ x_n)$，D'中的点是$(t_1,\ \cdots,\ t_n)$，变换T的表示式是

$$x_i=T_i(t_1,\ \cdots,\ t_n),\quad (t_1,\ \cdots,\ t_n)\in D'\quad (i=1,\ 2,\ \cdots,\ n).$$

则上面的式子就是

$$\int_D f(x_1,\ \cdots,\ x_n)\,dx_1\cdots dx_n=\int_{D'} f(T_1(t_1,\ \cdots,\ t_n)\cdots,$$

$$T_n(t_1,\ \cdots,t_n))\left|\frac{\partial(x_1,\ \cdots,x_n)}{\partial(t_1,\ \cdots,\ t_n)}\right|dt_1\cdots dt_n.$$

沙德定理

在逆映射的理论和积分的变量代换中，不等于0的雅可比行列式起了非常重要的作用。现在的问题是对给定的一个条件相当好的映射，例如它是连续可微的，那么它的雅可比行列式的零点有怎样的特点呢？

沙德(Sard)定理　设D是R^n内的一个开集，$f:D\to R^n$，并且f在D连续可微，令

$$E=\{x\in D|\det Df(x)=0\},$$

那么$f(E)$是一个零测度集。

证明　第一步：设A是D内的一个n维闭正方形，每一个边的长度都是l，对任意给定的$\varepsilon>0$，取一个充分大的自然数N，将A等分为N^n个n维闭正方形A_i $(i=1,\ 2,\ \cdots,\ N^n)$，A_i的边长是$\frac{l}{N}$，并且A_i和$A_j(i\neq j)$的内部不相交，同时对任何$x,\ y\in A_i$有

$$|Df(x)(y-x)-(f(y)-f(x))|<\varepsilon|y-x|\leqslant\varepsilon\sqrt{n}\frac{l}{N}.$$

由 2.3 的定理 1 这是办得到的。

第二步：如果 E 和某个 A_i 相交，固定 $x\in E\cap A_i$，有 $\det Df(x)=0$。让 $y=(y_1,\cdots,y_n)$ 在 A_i 内变动，考察

$$Df(x)(y-x)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x)\cdots\frac{\partial f_1}{\partial x_n}(x)\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\partial f_n}{\partial x_1}(x)\cdots\frac{\partial f_n}{\partial x_n}(x)\end{pmatrix}\begin{pmatrix}y_1-x_1\\ \cdots\\ y_n-x_n\end{pmatrix}$$

$$=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x)(y_1-x_1)+\cdots+\frac{\partial f_1}{\partial x_n}(x)(y_n-x_n)\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\partial f_n}{\partial x_1}(x)(y_1-x_1)+\cdots+\frac{\partial f_n}{\partial x_n}(x)(y_n-x_n)\end{pmatrix}.$$

记 $u_j=\frac{\partial f_j}{\partial x_1}(x)(y_1-x_1)+\cdots+\frac{\partial f_j}{\partial x_n}(x)(y_n-x_n)$。因为 $\det Df(x)=0$，所以当 $y=(y_1,\cdots,y_n)$ 在 A_i 内变动，点 $(u_1,u_2,\cdots,u_n)$ 在 R^n 内某一个 $n-1$ 维超平面 V 上，即 $\{Df(x)(y-x)\,|\,y\in A_i\}$ 在 V 上。

由第一步中给出的不等式可知道，固定 $x\in E\cap A_i$，$\{f(y)-f(x)\,|\,y\in A_i\}$ 和超平面 V 之间的距离小于 $\varepsilon\sqrt{n}\frac{l}{N}$，而 $f(x)$ 是一个常量，所以 $\{f(y)\,|\,y\in A_i\}$ 和 $n-1$ 维超平面 $V+f(x)$ 的距离小于 $\varepsilon\sqrt{n}\frac{l}{N}$。

第三步：让 x、y 都在 A_i 内变动，由 2.3 的定理 3 得到存在常数 $M>0$，使

$$|f(y)-f(x)|\leqslant M|y-x|\leqslant M\sqrt{n}\frac{l}{N}\quad(x,y\in A_i).$$

（在定理 3 的表述中，A_i 应该是一个 n 维的闭球，如今 A_i 却是一个 n 维闭正方形，但这是无关紧要的，从 2.3 的定理 3 的证明过程中容易看出对闭正方形仍成立。）

再由第二步的结论知道，如果 A_i 和 E 相交，则 $\{f(y)\,|\,y\in A_i\}$

含在一个 n 维柱体 C_i 内，C_i 的高是 $2\varepsilon\sqrt{n}\frac{l}{N}$ (这是因为 $f(y)$ 到超平面 $V+g(x)$ 的距离小于 $\varepsilon\sqrt{n}\frac{l}{N}$)，$C_i$ 的底是一个 $n-1$ 维球，半径是 $M\sqrt{n}\frac{l}{N}$ (这是因为对 A_i 内任何两点 x，y 都有 $|f(x)-f(y)|\leqslant M\sqrt{n}\frac{l}{N}$)，所以 C_i 的容积是 $2\varepsilon\sqrt{n}\frac{l}{N}\cdot\alpha\left(\frac{l}{N}\right)^{n-1}=2\alpha\times\sqrt{n}\left(\frac{l}{N}\right)^n\varepsilon$ (其中 $\alpha>0$ 是一个常数)。

第三步的结论是：当 A_i 和 E 相交，$f(A_i)$ 将被柱体 C_i 覆盖，C_i 的容积是 $2\alpha\sqrt{n}\left(\frac{l}{N}\right)^n\varepsilon$。

第四步：估计 $f(A\cap E)$。

因为 $f(A\cap E)$ 至多被 N^n 个上述那种柱体 C_i 所覆盖，这些 C_i 的容积之和不超过

$$N^n\cdot2\alpha\sqrt{n}\left(\frac{l}{N}\right)^n\varepsilon=2\alpha\sqrt{n}\,l^n\varepsilon.$$

所以 $f(A\cap E)$ 是一个零容度集(注：在零容度集的定义中，是利用 n 维闭矩形作覆盖的，但这里却是用 n 维柱体作覆盖的。为什么可以这样做?事实上两者无本质差别)。

到现在为止，证明了对 D 内任何一个闭矩形 A，$f(A\cap E)$ 是零容度集，$f(E)$ 如何呢?

最后一步：利用下面的命题：R^n 中的任何一个开集 D 可以表示为可列无限多个 n 维闭矩形 A_1，A_2，…，A_n，…之并，即 $D=\bigcup_{n=1}^{\infty}A_n$，并且 A_i 和 $A_j(i\neq j)$ 的内部不相交 (见第 61 页引理)，由前四步知道，对每一个 A_n，$f(A_n\cap E)$ 是一个零容度集。而它至多是可列无限多个，因此 $f(E)$ 是一个零测度集。证毕。

习　题

1. 利用定理 6 证明：若 f、g 都在闭矩形 A 可积，则 $f\cdot g$ 也在 A 可积。
2. 利用定理 6 证明：若 f_1、f_2、…，f_n 都在闭矩形 A 可积，令

$$F(x)=\max_{n}\{f_1(x),\cdots,f_n(x)\};$$

$$G(x)=\min_{n}\{f_1(x),\cdots,f_n(x)\}\quad(x\in A),$$

则 F、G 都在 A 可积。

3. 设 f 在闭矩形 A 可积，并且 $m\leqslant f(x)\leqslant M(x\in A)$．又设一元函数 φ 在 $[m,M]$ 连续，证明 $\varphi\circ f$ 在 A 可积。如果将 φ 的“连续”改为“可积”，问 $\varphi\circ f$ 是否在 A 可积？

4. 设 $f_1,f_2,\cdots,f_n,\cdots$ 是定义在 $[a,b]$ 上的一列可积函数，并且 $\lim\limits_{n\to\infty}f_n(x)=f(x)\ (x\in[a,b])$ 存在，问 f 是否也在 $[a,b]$ 可积？

第4章 微分流形

§4.1 拓扑空间

在欧几里得空间上由于有了欧几里得范数，从而产生了欧几里得距离，利用距离就可以引进开集、闭集、收敛和连续等一系列极重要的概念。收敛和连续都是用距离来表达的，然而也可以换一个方式用开集来表达，例如 $x_n \to x$ 就是指对含有点 x 的任意一个开集 U，当 n 充分大时 $x_n \in U$。又如映射 f 连续是指对任何开集 V，其逆象 $f^{-1}(V)$ 是开集，这种表达方式完全脱离了范数或距离，它似乎表明完全可以把开集作为出发点来建立整个数学分析。不过，在欧几里得空间中，开集是利用范数来定义的，出发点是范数，但在一个一般的空间中，怎样合理的规定何种集合是开集，而这种规定不必借用范数。回忆一下在欧几里得空间中开集有哪些最重要的特性？仔细分析就会发现凡是和开集有关的一切概念、定理和论证都只用到下面三条最基本的特性：

(i) 任意个（有限，可列无限或不可列无限多个）开集的并集仍旧是开集。

(ii) 有限个开集的交集仍旧是开集。

(iii) 整个空间 R^n 是开集，空集 $\varnothing$ 是开集（后者也可以认为是规定）。

这启发我们，在一个空间内可以不必引进范数或距离，只要引进满足上述三条的开集，则就可以建立内点，外点，边界点，闭集，紧集，聚点，收敛，连续等一系列基本概念和基本理论，事实正是如此。

拓扑

设 X 是一个集合，τ 是 X 中某些子集组成的集类，如果 τ 满足

下面三条性质(或者说满足下面三条公理)：

(i) τ 内任意个集合的并集仍旧属于 τ；

(ii) τ 内有限个集合的交集仍旧属于 τ；

(iii) X 属于 τ，空集 $\varnothing$ 属于 τ。

就称 τ 是 X 上的一个**拓扑**，又称 τ 中的集是 X 内的**开集**。设点 $x\in X$，称含有 x 的开集是 x 的一个**邻域**。当 X 装备了拓扑结构 τ 之后，称 (X, τ) 是**拓扑空间**，有时简单地记之为 X。换句话说，所谓拓扑空间 (X, τ) 就是指在集合 X 上建立了满足上述公理的开集。从直观上看，建立了开集意味着给出了"靠近"的概念，X 内的两个点 a、b，如果同在某个开集内，就表明 a、b 之间有一种靠近的程度，这个程度是用它们同属于某个开集来表征的。

例 1 设 $X=\{a, b, c\}$，它是由三个元素 a、b、c 组成的。又设

$$\tau=\{\{a\}, \{a, b\}, X, \varnothing\},$$

容易验证 τ 是一个拓扑，(X, τ) 是一个拓扑空间。再设

$$\tau'=\{\{a\}, \{a, b\}, \{a, c\}, X, \varnothing\}.$$

τ' 也是一个拓扑，(X, τ') 也是一个拓扑空间。由于 (X, τ) 和 (X, τ') 装备着不同的拓扑结构，所以它们是不同的拓扑空间。

例 2 在实数集 X 上装备以下拓扑 τ：由所有开区间、任意个开区间的并以及空集 $\varnothing$ 组成。这个拓扑结构中的开集就是通常数学分析中由距离产生的开集，称 τ 是**通常的拓扑**。

在实数集 X 上还可以建立其他的拓扑，例如 τ'：由所有包含 0 点的开区间和 $\varnothing$ 组成。不难验证 τ' 是拓扑。

(X, τ) 和 (X, τ') 都是拓扑空间，但它们是两个不同的拓扑空间。

例 3 设 $(X, \|\ \|)$ 是一个赋范线性空间，$\|\ \|$ 是范数，对任何一点 $a\in X$ 和任意一个实数 $\delta>0$，记

$$O(a, \delta)=\{x\in X \mid \|x-a\|<\delta\}.$$

它是以点 a 为中心，以 δ 为半径的开球。设 E 是 X 的一个子集，如果在 E 内的每个点 y 都存在一个开球 $O(y, \eta)\subset E$，其中 $\eta>0$ 可能与点 y 有关，则称 E 是 X 内的一个开集。令 τ 是 X 内的所有开集以及 X 自身和空集 $\varnothing$ 组成的集类，容易验证 τ 是一个拓扑，称它是由范数产生的拓扑。因此任何一个赋范线性空间一定是一个拓扑空间，其拓扑由范数产生。

例 4 （两个极端的情形） 设 X 是一个集合，

τ_1：由 X 的一切子集组成；

τ_2：由 X 和 $\varnothing$ 组成（仅两个元素）。

τ_1 和 τ_2 都是拓扑，称 (X, τ_1) 是离散的拓扑空间，τ_1 是离散拓扑。在离散的拓扑空间中，由一个点 $x\in X$ 组成的集 $\{x\}$ 是 X 中的一个开集。（为什么称 (X, τ_1) 是"离散"的，在后面要说明。）称 (X, τ_2) 是平凡的拓扑空间（或最粗糙的拓扑空间）。

例 1 中的两个拓扑 τ 和 τ' 虽然不同，但它们却有下列关系：τ 中的开集（也就是 τ 的元素）都在 τ' 中，即 $\tau\subset\tau'$，这意味着 τ' 中的开集更多一些，则称拓扑 τ' 强于拓扑 τ，或者说拓扑 τ 弱于拓扑 τ'。一般说来，在同一个集合上装备了两个不同的拓扑 τ 和 τ'，如果 $\tau\subset\tau'$，就称 τ' 强于 τ（或 τ 弱于 τ'）。但是不要以为同一个集合上的任意两个不同的拓扑结构都可以比较强弱，例如在例 1 中又设 $\tau''=\{\{c\}, \{b, c\}, X, \varnothing\}$，$\tau''$ 也是一个拓扑，但是它和 τ，和 τ' 都不可比较。又如在例 2 中设 τ'' 是所有包含闭区间 $[1, 2]$ 的开区间以及 $\varnothing$ 组成的集类，τ'' 是一个拓扑，但它和 τ' 不可比较。

收敛

现在将欧几里得空间中收敛的概念和一般拓扑空间中的收敛的概念作一个对比，叙述如下：

欧几里得空间 R^n	拓扑空间 (X, τ)
设点列 $\{x_n\}\subset R^n$，点 $x\in R^n$，如果对以 x 为中心的任何一个开球 O，存在 N，当 $n>N$ 时，有 $x_n\in O$。则称 $\{x_n\}$ 收敛，x_n 收敛于 x，记为 $x_n\to x$。	设点列 $\{x_n\}\subset X$，点 $x\in X$，如果对含有 x 的任何一个开集 O，存在 N，当 $n>N$ 时，有 $x_n\in O$。则称 $\{x_n\}$ 收敛，x_n 收敛于 x，记为 $x_n\to x$。

以上两者的表达方式几乎是一样的。

要注意的是：由于拓扑空间是非常一般的，在这个非常一般的空间中只有开集满足的三条性质可以作为出发点加以运用，因此它不象欧几里得空间那么细腻，例如在欧几里得空间中任何两个不同的点 x 和 y，可以作两个分别以 x 和 y 为中心的适当小的开球，使得这两个开球不相交，换句话说，可以用不相交的开集把不同的两点 x 和 y 隔离开，但在一般的拓扑空间中，仅仅从开集的三条性质出发是推断不出这一隔离性质的（附带说一下，一般的赋范线性空间是具有这一性质的。为什么？是很容易证明的），正由于一般的拓扑空间不象欧几里得空间那么细致，所以它将会产生许多奇怪的"不正常"的现象。

例 5 设 X 是实数集，τ_1 是 X 上通常的拓扑（见例 2），τ_2 是由所有包含 $(0, 1)$ 的开区间以及空集 $\varnothing$ 组成，则 (X, τ_1) 和 (X, τ_2) 都是拓扑空间。考察数列

$$x_n=\frac{1}{n},\quad y_n=c+\frac{1}{n}\ (c>0\text{ 是一个常数}),$$

$$z_n=\begin{cases}\dfrac{1}{n}, & n\text{ 是奇数；}\\ 1-\dfrac{1}{n}, & n\text{ 是偶数。}\end{cases}\qquad (n=1, 2, \cdots)$$

在 (X, τ_1) 内 $x_n\to 0$，$y_n\to c$，$\{z_n\}$ 不收敛。但在 (X, τ_2) 中情况大为两样，$\{x_n\}$ 收敛于任何实数；而 $\{y_n\}$ 收敛于任何不小于 c 的实数，$\{z_n\}$ 也收敛于任何实数。这是因为对任何实数 α，含有 α 的任何开

集必含有$(0, 1)$，从而含有 x_n 和 $z_n(n=1, 2, \cdots)$，所以$\{x_n\}$和$\{z_n\}$都收敛于α。另外对任何 $\beta\geqslant c$，含有β的任何开集O必含有 $(c, \beta]$，而$\beta\geqslant c$，所以存在N，当$n>N$时，$y_n\in O$，即 $y_n\to\beta$。

由于拓扑τ_1强于τ_2，因此不存在(X, τ_1)内收敛但在(X, τ_2)中不收敛的数列。

例 6 设X是实数集，拓扑τ是由所有左开右闭的区间$(-a, 0]$ ($a>0$，是任何一个实数)以及X，$\varnothing$组成。在(X, τ)内数列$\left\{\frac{1}{n}\right\}$不收敛于0，但收敛于任何$c>0$。

例 7 设X是任一集，τ是X上的离散拓扑(见例4)，对任何点$x\in X$，因为由单点x组成的集$\{x\}$是一个开集，它含有x，并且在这个开集内除x外不再有其他的点，所以在离散的拓扑空间内，只有点列$x, x, \cdots, x, \cdots$收敛于x。如果把收敛看作"越来越靠近"的话，那么，在离散的拓扑空间(X, τ)中只有$x, x, \cdots, x, \cdots$越来越靠近x，由其他非x的点组成的点列都不会越来越靠近x，点x成了"孤立点"，整个空间是离散的，这就是离散的拓扑空间的直观解释。

例 8 在最粗糙的拓扑空间中，任何点列都收敛，并且收敛于空间中的任一点。

出现这种"不正常"的现象是因为我们只能从开集的三条性质出发，它不够细腻。怎样才能防止这种现象出现呢？在本节最后要扼要的讨论它。

闭集、紧集等概念

和欧几里得空间相仿，可以引进内点、外点、边界、闭集、紧集等一系列重要的概念。

设(X, τ)是一个拓扑空间，E是X的一个子集。

(i) 内点　设$x\in E$，如果存在一个含有x的开集$O\subset E$，就称x是E的一个内点。显然，开集U内的每一点都是U的内点。

(ii) 外点　设$x\in X$，如果存在一个含有x的开集O，$O\subset X-$

E（或者写为 $O\cap E=\emptyset$），就称 x 是 E 的一个外点。显然，$x\notin E$。

(iii) 边界点　设 $x\in X$ (x 可以属于 E，也可以不属于 E)，如果对含有 x 的任何开集 O，O 中既有 E 的点又有非 E 的点，即 $O\cap E\neq\emptyset$，$O\cap(X-E)\neq\emptyset$，就称 x 是 E 的一个边界点，E 的所有边界点組成 E 的边界。

(iv) 聚点　设 $x\in X$ (x 可以属于 E，也可以不属于 E)，如果对含有 x 的任何开集 O，$O-\{x\}$ 中必有 E 的点，即 $(O-\{x\})\cap E\neq\emptyset$，就称 x 是 E 的一个聚点。要注意的是，E 的聚点可能不属于 E。

(v) 闭集　设 E 是 X 的一个子集，如果 E 的补集 $X-E$ 是 X 中的开集，就称 E 是闭集。也就是说，开集的补集是闭集，闭集的补集是开集。X 和空集 $\emptyset$ 既是开集又是闭集。

(vi) 闭包　记 $\bar{E}=E\cup\{E$ 的所有聚点$\}$，称 $\bar{E}$ 是 E 的闭包。例如实数集上带有通常的拓扑，$E=(a,b)$，则 $\bar{E}=[a,b]$。

有的书上关于闭集的定义是：设 E 是 X 的一个子集，如果 E 的每一个聚点（假定它存在的话）都属于 E，即 $E=\bar{E}$，就称 E 是闭集，这个定义和(v)中的定义是等价的，这就是下面的性质：

性质　设 $E\subset X$，则下列两个命题是等价的：

(1) E 的补集 $X-E$ 是开集；

(2) $E=\bar{E}$。

（因此它们中的任何一个都可以作为闭集的定义。）

证明　设 $X-E$ 是开集，又设 $x\notin E$ 即 $x\in X-E$，这表明存在一个含有 x 的开集 O，其中不含有 E 的点，即 x 不是 E 的聚点。换句话说，如果 x 是 E 的聚点，则 x 必属于 E。

反过来，设 $E=\bar{E}$。又设 $x\notin E$ 即 $x\in X-E$。因为 E 的所有聚点都在 E 内，所以 x 不是 E 的聚点，即存在某个含有 x 的开集 O_x，$O_x\cap E=\emptyset$ 即 $O_x\subset X-E$，而 $X-E=\bigcup\limits_{x\in X-E}O_x$，右端是任意个开集之并，所以 $X-E$ 是开集。证毕。

(vii) 紧集　设 $\{O_\alpha\}$ 是 E 的任意一个开覆盖（即每一个 O_α 是

开集，并且 $E\subset\bigcup_\alpha O_\alpha$)，如果在$\{O_\alpha\}$中总可以选出有限个开集覆盖$E$，就称$E$是一个紧集。

例 9 在欧几里得空间(或者任意一个赋范线性空间)内，设x_0是E的一个聚点，则必存在一个点列$\{x_n\}\subset E$，$x_n\neq x_0(n=1, 2, \cdots,)$，$x_n\to x_0$。但在一般的拓扑空间$(X, \tau)$内，这一事实不一定成立。例如设$X$是任意一个不可列集，$O$是$X$的一个子集，如果$O$的补集$X-O$是可列集或有限集，就说$O$具有可列补。设$\tau$是由所有具有可列补的子集$O$以及空集$\varnothing$组成的集类，不难验证$\tau$是一个拓扑。称这一拓扑是可列补拓扑。

在可列补拓扑空间内，设$x_0\in X$，考察$E=X-\{x_0\}$，显然每个含x_0的开集必含有E的点，所以x_0是E的一个聚点，是否存在一个点列$\{x_n\}\subset E$，$x_n\neq x_0\ (n=1, 2, \cdots)$，$x_n\to x_0$呢？在$E$中任意取一个点列$\{x_n\}$，$x_n\neq x_0(n=1, 2, \cdots)$，现在作

$$O=X-\{x_1, x_2, \cdots, x_n, \cdots\},$$

因为$X-O=\{x_1, x_2, \cdots, x_n, \cdots\}$是可列集，所以$O$是含$x_0$的一个开集，但这个开集中却不含$\{x_n\}$，所以$x_n$不收敛于$x_0$，即$E$内任何点列(除$x_0, x_0, \cdots, x_0, \cdots$ 外)都不会收敛于x_0。

拓扑子空间

设(X, τ)是一个拓扑空间，S是X的一个子集，要在S上建立拓扑τ_s，使得τ_s和原来的拓扑τ相容，这就是说，如果点列$\{x_n\}\subset S$，点$x\in S$，在(X, τ)内$x_n\to x$和在(S, τ_s)内$x_n\to x$是相同的，如何建立τ_s？

设O是X内的开集，记

$$O_s=O\cap S,$$

把O_s作为S内的一个开集。对X内的所有开集O都作出相应的O_s，这些O_s的全体记为τ_s，即

$$\tau_s=\{O\cap S\,|\,O\in\tau\},$$

τ_s就是所要求的拓扑。称τ_s是τ在S上的诱导拓扑，(S, τ_s)是

(X, τ)的拓扑子空间。

例 10 设X是实数集，τ是通常的拓扑，$S=[0,1)$，由τ诱导出的拓扑是

$$\tau_s=\{O\cap[0,1)\mid O\in\tau\}.$$

例如$\left[0,\frac{1}{3}\right)=\left(-1,\frac{1}{3}\right)\cap[0,1]$是$S$内的开集，从而$\left[\frac{1}{3},1\right)$是$S$内的闭集。

连续和同胚

设(X, τ_x)和(Y, τ_y)是两个拓扑空间，它们可以是相同的也可以是不同的。如果对每一点$x\in X$，通过关系f在Y内存在唯一的一点y与这个x对应，就称f是从X到Y的映射，记为$f: X\to Y$。称y是x在f作用下的象，x是y的一个逆象，记为$x\mapsto y=f(x)$。有时为了表明f是两个拓扑空间之间的映射，又记$f:(X, \tau_x)\to(Y, \tau_y)$。此外，和欧几里得空间上的映射相仿地可以引进单射、满射和双射的概念，这里不一一叙述了。

在欧几里得空间中，可以用"开集的语言"来定义映射的连续性(参见1.3定理2)，而不必借用"距离"，现在拓扑空间内已经装备了开集，因此相应的连续概念也就不难产生了。

设$f:(X, \tau_x)\to(Y, \tau_y)$，如果对$Y$中任意一个开集$V$，它的逆象

$$U=f^{-1}(V)=\{x\in X\mid f(x)\in V\}$$

是X中的开集，就称f是从X到Y的一个连续映射。当X和Y都是欧几里得空间(带有通常的拓扑)时，上述定义就是欧几里得空间上连续映射的定义。映射的连续性和空间的拓扑结构有关，在某一个拓扑结构下是连续的映射，换成另一个拓扑结构之后可能变得不连续了。

例 11 设X是实数集，映射$f: X\to X$的定义是$f(x)=\sin x$ $(x\in X)$。f是否连续？这要看X上的拓扑结构了。设τ_1是X上的通常拓扑；τ_2是由所有包含0的开区间以及空集$\varnothing$组成的拓扑，

则 $f:(X, \tau_1)\to(X, \tau_1)$ 是连续的，但 $f:(X, \tau_2)\to(X, \tau_1)$ 是不连续的。后者是因为在 τ_1 中存在开集 $\left(\frac{1}{2}, \frac{2}{3}\right)$，其逆象 $U=\left\{x\in X \,\middle|\, \frac{1}{2}<\sin x<\frac{2}{3}\right\}$ 不是含 0 的开区间，即 U 不是 τ_2 中的开集。

例 12 设 (X, τ_x) 和 (Y, τ_y) 是两个拓扑空间，$f:X\to Y$。如果对 X 中的每一点 x 以及 X 中的任何一个收敛于 x 的点列 $\{x_n\}$ 在 Y 内都有 $f(x_n)\to f(x)$，f 是否连续？在欧几里得空间（带有通常拓扑）内，回答是肯定的。但在一般拓扑空间中却未必如此。例如设 X 是实数集，τ_1 是 X 上的通常拓扑，τ_2 是可列补拓扑（例 9），考察映射

$$g:(X, \tau_2)\mapsto(X, \tau_1)$$

$$x\mapsto x^2.$$

在 (X, τ_2) 中，对每一个 x，只有 $x, x, \cdots, x, \cdots$ 收敛于 x，其象 $x^2, x^2, \cdots, x^2, \cdots$ 当然也收敛于 x 的象 x^2，但 τ_1 中的开集 $(1, 2)$ 其逆象却不是 τ_2 中的开集，因此 g 不连续。

现在引进拓扑学中一个重要的概念，它在下面的课文中将起关键作用。设 (X, τ_x) 和 (Y, τ_y) 是两个拓扑空间，f 是从 X 到 Y 的双射（因此逆映射 f^{-1} 存在），如果 f 又是从 X 到 Y 的连续映射，并且 f^{-1} 是从 Y 到 X 的连续映射，就称 f 是同胚映射。当两个拓扑空间之间存在一个同胚映射时，就称这两个空间同胚。

如果两个空间是同胚的，则这两个空间之间不仅点与点一一对应，而且开集与开集之间也一一对应，这表明它们有相同的拓扑结构，从拓扑学的观点看，可以把同胚的两个拓扑空间看作是相同的，这如同在代数学中所说的"同构的两个空间是相同的"一样。

再从直观上看，如果把拓扑空间看作一块有弹性的橡皮薄膜，那么同胚映射就是将这块薄膜作拉伸压缩或弯曲，但不准撕开或粘贴，薄膜经同胚映射后，形状可能改变了，但两者的点与点之间，开集与开集之间是一一对应的。

例 13 设 X 是实数集，带有通常的拓扑，$(-1, 1)$ 是 X 的拓扑

子空间。作映射 f 是

$$f(x)=\mathrm{tg}\frac{\pi}{2}x, \quad x\in(-1,1)。$$

很明显，f 是从 $(-1,1)$ 到 X 的同胚映射，所以 X 和 $(-1,1)$ 同胚。

例 14 在 3 维欧几里得空间 R^3(带有通常拓扑)中，上半球面 $S=\{(x,y,z)\in R^3|x^2+y^2+z^2=1, z>0\}$ 和开圆盘 $D=\{(x,y,z)\in R^3|z=0, x^2+y^2<1\}$ 同胚（注：S 和 D 都看成 R^3 的拓扑子空间）。这是因为设映射 $f:D\to S$ 是

$$x=x, \quad y=y, \quad z=\sqrt{1-x^2-y^2}, \quad (x,y)\in D,$$

f 是从 D 到 S 的同胚映射，所以 D 和 S 同胚。

例 15 设 X 是实数集，τ_1 是 X 上的通常拓扑，τ_2 是由所有含 0 的开区间以及空集 $\varnothing$ 组成的拓扑，则 (X,τ_1) 和 (X,τ_2) 不同胚。

可数公理和隔离公理

这里，简要地介绍一下怎样防止拓扑空间过于一般化而产生的许多很不细腻也很不正常的现象。作为一个拓扑空间，它只装备着满足三条性质的开集体系，其余的一切只能从这三条性质出发，当然会显得不够细致了，为了使它更细致些，就必须另外加进一些公理，在这些公理的限制下，空间的性质将会变得更好一些，这些公理共有两套，一套是可数公理，另一套是隔离公理。

可数公理说的是开集的“数量”。如果拓扑空间中的开集有不可列无限多个，问能不能从中选出可列个就够用了，这就是可数公理的作用，它共有两个，如下所述：

第一可数公理 如果拓扑空间 (X,τ) 内的每一点 x，存在可列个开集 $O_n(n=1,2,\cdots)$，使对任何一个含有 x 的开集 V，$\{O_n\}$ 中存在一个 $O_{n_0}\subset V$，就称 (X,τ) 满足第一可数公理(或 A_1 公理)。直观地说，在满足第一可数公理的空间中，任意固定一点 x，它的邻域(即含有 x 的开集)可能会有不可列无限多个，但总可以选出

可列无限多个就足以和那不可列无限多个相当。

第二可数公理 设(X, τ)是拓扑空间，如果存在可列个开集$O_n(n=1, 2, \cdots)$，使得任何一个开集都可以表示为$\{O_n\}$中某些开集的并，就称(X, τ)满足第二可数公理（或A_2公理）。直观地说，在满足第二可数公理的空间中，其开集可能是不可列无限多个，但总可以选出可列无限多个就足够了。例如欧几里得空间R^n（带有通常拓扑），就是满足第二可数公理的空间（请读者证明）。

隔离公理说的是点和点之间，或者点和闭集之间，闭集和闭集之间的一种隔离性质。隔离公理有好几个。例如：

豪斯多夫(Hausdorff)公理 设(X, τ)是拓扑空间，如果对X内任何两点$x, y(x \neq y)$，总存在两个开集O_x和O_y，$O_x \cap O_y = \varnothing$，并且$x \in O_x$，$y \in O_y$。就称$(X, \tau)$满足豪斯多夫公理（或$T_2$公理），又称$(X, \tau)$是一个*豪斯多夫空间*。例如任何一个赋范线性空间一定是豪斯多夫空间。带有可列补拓扑的空间（例9）不是豪斯多夫空间。在豪斯多夫空间内如果点列x_n收敛，则它只收敛于唯一的一点。这些都作为习题，请读者自证。

除了T_2公理外，还有其他的隔离公理，如T_0，T_1，T_3，T_4等，这里不一一叙述了。

在拓扑空间上加进可数公理和隔离公理的限制以后，拓扑空间的性质将有所改善。在怎样的公理下会改善到何种程度，有兴趣的读者可以参看任何一本有关点集拓扑的书，这里不作介绍了。

拓扑空间的度量化

在一个赋范线性空间$(X, \| \ \|)$上，由范数$\| \ \|$可以产生开球，从而产生开集，因此它是一个拓扑空间。反过来，如果在一个拓扑空间(X, τ)上可以建立某个范数$\| \ \|$，使得由$\| \ \|$产生的所有开集就是τ，那么就称拓扑空间(X, τ)是*可度量化的*。任何一个拓扑空间是否都可度量化呢？答案是否定的。

例 16 设$X=\{a, b\}$，$\tau=\{\{a\}, \varnothing, X\}$，$(X, \tau)$是拓

扑空间。如果它可度量化，设范数是$\|\ \|$，又设$\|a-b\|=r>0$，作开球

$$O\left(b,\frac{r}{2}\right)=\left\{x\in X\,\middle|\,\|x-b\|<\frac{r}{2}\right\}=\{b\},$$

即$\{b\}$是一个开集，但$\{b\}\notin\tau$，这就导致矛盾。

确实还存在许多在理论上和应用上都很有价值的拓扑空间是不可度量化的，但已超出本书的范围，这里不作介绍了。

习　题

1．设A、B都是某拓扑空间的子集，试证明：

(i) $A\subset\bar{A}$;　(ii) $\bar{\bar{A}}=\bar{A}$;　(iii) $\overline{A\cup B}=\bar{A}\cup\bar{B}$.

又：$\overline{A\cup B}$是否等于$\bar{A}\cap\bar{B}$? $\bar{A}-\bar{B}$是否等于$\overline{A-B}$? 如果不相等，它们之间有什么关系?

2．设(X,τ)是拓扑空间，$A\subset X$，如果$\bar{A}=X$，就称A在X内是稠密的。设X是实数集，装备着下列三种拓扑：τ_1是通常拓扑；τ_2是可列补拓扑；τ_3是所有含点0的开区间以及空集组成的拓扑。再设Q是所有有理数组成的集。问Q分别在(X,τ_1)，(X,τ_2)，(X,τ_3)内稠密吗?

3．设(X,τ_1)和(Y,τ_2)是两个拓扑空间，$f:X\to Y$连续，证明对X中任何紧集K，$f(K)$是Y中的紧集。

4．设(X,τ_1)和(Y,τ_2)都是拓扑空间，$f:X\to Y$，试证明下列命题是等价的：

(i) f在X连续；

(ii) Y中的每一个闭集C，$f^{-1}(C)$是X中的闭集；

(iii) 对X的每一个子集A，$f(\bar{A})\subset\overline{f(A)}$。

5．设(X,τ)是拓扑空间，R是1维欧几里得空间，$f:X\to R$连续，K是X中的一个紧集，试证明f在K上有最大值和最小值。

6．设

$$1:(X,\tau_1)\to(X,\tau_2)$$
$$x\mapsto x,$$

称1是从X到X的恒等映射，它连续吗?为了使1连续，问τ_1和τ_2应有什么关系?并给出1连续的充分和必要条件。

7．证明欧几里得空间满足A_2公理和T_2公理。

8．证明赋范线性空间满足A_1公理和T_2公理。还满足A_2公理吗?

9．证明在豪斯多夫空间中，如果$x_n\to x$，$x_n\to y$，则$x=y$。

§4.2 微分流形

在3维欧几里得空间 R^3 中，我们可以讨论1维光滑曲线、2维光滑曲面；讨论曲线的切线和曲面的切平面；讨论曲线积分和曲面积分以及它们之间的关系。但还有不少问题值得进一步探索，例如：(i)为什么要在曲线的前面冠以"1维"，在曲面的前面冠以"2维"？从解析几何和线性代数的观点看，只有通过原点的直线是1维的，通过原点的平面是2维的，那么1维曲线，2维曲面又如何解释呢？(ii)更重要的是，在理论和应用的研究中，并不限于3维或3维以下的欧几里得空间，那么在更高维的空间中如何讨论光滑的"几何体"？如何正确表达和研究这些光滑的"几何体"？怎样把 R^3 中的光滑曲面的概念拓广到更高维的空间甚至更一般的拓扑空间中？(iii)怎样把曲线的切线和曲面的切平面的概念拓广到这些光滑的"几何体"上？(iv)如何定义光滑的"几何体"上的积分？使得这一积分确实是曲线积分和曲面积分的自然的拓广。又例如在 R^3 中有三个著名的公式将重积分、曲面积分、曲线积分两两联系起来，它们是格林(Green)公式、高斯(Gauss)公式和斯托克斯(Stokes)公式，那么在更高维的"几何体"上定义的积分是否也有相应的公式？这些问题构成本章和以后几章的中心，我们将逐步展开研究之。

流形

先考察两个例题：

例1 4.1的例14给出了 R^3 中的上半单位球面 $S=\{(x, y, z)\in R^3 | x^2+y^2+z^2=1, z>0\}$ 和开圆盘 $D=\{(x, y, 0)\in R^3 | x^2+y^2<1\}$ 同胚，而开圆盘 D 显然又和整个2维欧几里得空间同胚，因此 S 和 R^2 同胚，从拓扑学的观点上看，S 和 R^2 是一样的，在这个意义上，我们说 S 是一个2维曲面。

例2 设 S 是 R^3 中一个球面，$S=\{(x, y, z)\in R^3 | x^2+y^2+z^2$

$=a^2\}$，在 S 上装备着由 R^3 中通常拓扑诱导出来的拓扑。S 和例1中的半球面不同，它不可能和 R^2 中任何开集同胚，又如何说它是2维的呢？

这只要作一次"手术"就可以了。将 S 剖成两片 S_1 和 S_2，S_1 是上半球面 $\{(x, y, z)\in R^3|x^2+y^2+z^2=a^2, z>0\}$，$S_2$ 是下半球罐 $\left\{(x, y, z)\in R^3|x^2+y^2+z^2=a^2, z<\frac{a^2}{2}\right\}$（图 4-1）。$S_1$ 和 S_2 上的拓扑都是由 R^3 内的通常拓扑诱导出来的。并且 S_1 和 S_2 都是开集，它们覆盖了 S。

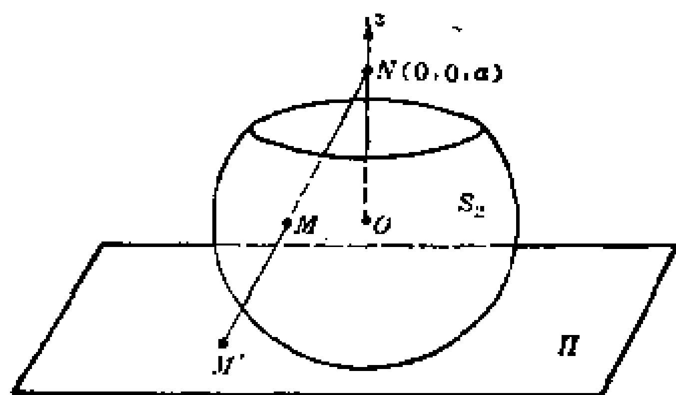


图 4-1

由例1知道 S_1 和平面上的一个开圆盘同胚。在图4-1中，点 $N=(0, 0, a)$ 是球面 S 的北极，平面 Π 是球面 S 在南极 $(0, 0, -a)$ 的切平面。通过点 N 作射线穿过球罐 S_2 和平面 Π，设交点分别是 M 和 M'，作映射 $\varphi: M\mapsto M'$，在这一映射下容易看出 S_2 和平面上的某个开圆盘同胚。

可见 S 虽然不和平面上的任何一个开集同胚，但它有两个开集 S_1 和 S_2 组成的开覆盖，每一个开集都和平面上的某个开圆盘同胚，换句话说，从整体说来，S 不和 R^2 内的任何开集同胚，但 S 的每一个局部（开集）都和 R^2 内某个开集同胚，所以我们也称 S 是一个2维曲面。

将它略加抽象，就得到流形的概念。

设 M 是一个拓扑空间，并且是豪斯多夫空间，$\{U_\alpha\}$ 是 M 的开

覆盖。如果对每一个开集 U_α，联系着一个映射 $\varphi_\alpha: U_\alpha \to V_\alpha$，$V_\alpha$ 是 n 维欧几里得空间内的一个非空开集，也可以不妨假设它是 n 维开球或 n 维开矩形(图 4-2)，并且 φ_α 是同胚映射，即 U_α 和 V_α 同胚，就称M是一个 n 维流形。从流形的定义中可使我们知道两件事：

(i) M虽然不是欧几里得空间，但是它可以"局部欧几里得空间化"，M的每一个局部 U_α 都和 R^n 中的一个开集 V_α 同胚，从拓扑学的观点看，U_α 和 V_α 相同。

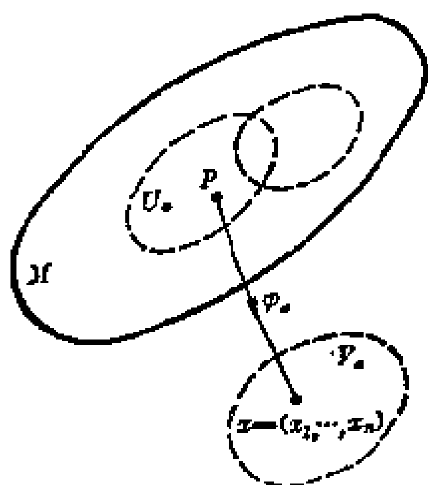


图 4-2

(ii) M内的点并没有坐标，但可以安装"局部坐标"，由于 U_α 和V_α同胚，它们的点与点一一对应，设点 $p\in U_\alpha$，通过同胚映射 φ_α，点 p 映射为 V_α 中的点 x，设 $x=(x_1, \cdots, x_n)$，就把 $(x_1, \cdots, x_n)$看成点 p 的局部坐标，并用$(x_1, \cdots, x_n)$来表示点 p (图 4-2)。

从直观上概括地说：所谓 n 维流形就是可以局部 (n 维)欧几里得空间化的一个拓扑空间，其中的点可以安装局部坐标。

微分流形

在 n 维流形上加上"微分结构"就成为微分流形。

设M是一个 n 维流形，这表明M有一个开覆盖$\{U_\alpha\}$，并且每一个 U_α 联系着一个映射 $\varphi_\alpha: U_\alpha \to V_\alpha$，$\varphi_\alpha$ 是同胚映射。如果 U_α 和 U_β 都在$\{U_\alpha\}$内，并且 $U_\alpha \cap U_\beta \neq \varnothing$，设 $p\in U_\alpha\cap U_\beta$，在 φ_α 的

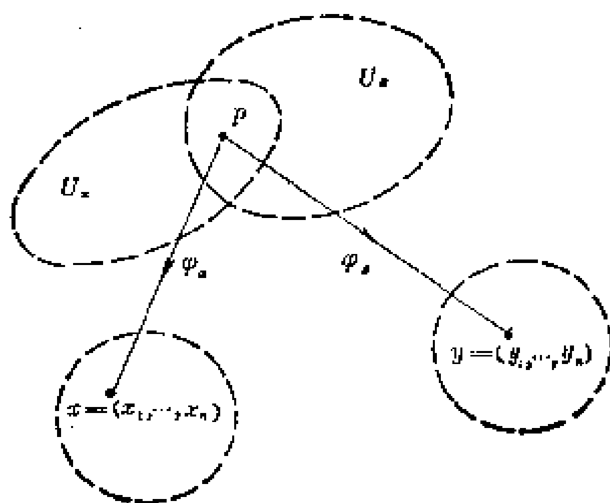


图 4-3

作用下将点 p 映射为 $x=(x_1, \cdots, x_n)$，在 φ_β 的作用下将点 p 映射为 $y=(y_1, \cdots, y_n)$(图 4-3)。

同一个点 p，既可以用坐标 $x=(x_1, \cdots, x_n)$ 表示，又可以用坐标 $y=(y_1, \cdots, y_n)$ 表示，于是这两个坐标之间有一个坐标变换式，即

$$\varphi_\beta \circ \varphi_\alpha^{-1}: \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$
$$x=(x_1, \cdots, x_n) \mapsto y=(y_1, \cdots, y_n),$$

映射 $\varphi_\beta \circ \varphi_\alpha^{-1}$ 是从 $\varphi_\alpha(U_\alpha \cap U_\beta)$ 到 $\varphi_\beta(U_\alpha \cap U_\beta)$ 的双射，并且又是 n 维欧几里得空间内的映射。如果它可微并且它的逆映射也可微，就称 M 是 n 维微分流形。又称 U_α 是坐标邻域，φ_α 是坐标映射，$(U_\alpha, \varphi_\alpha)$ 是坐标图，$\{(U_\alpha, \varphi_\alpha)\}$ 是 M 的坐标图册。这好比用一本地图册来表示地球表面上的地理疆界一样，地球表面（近似地）是球面，它不和平面上的任何开矩形同胚，但可以把整个地球面用有限个开集来覆盖，每一个开集同胚于地图册上的一页图，后者就是平面上的一个开矩形，这样一本地图册就足以表现了全世界的地理疆界。

如果 M 是一个 n 维微分流形，对一切 $U_\alpha \cap U_\beta \neq \varnothing$ 的 α 和 β，可微映射 $\varphi_\beta \circ \varphi_\alpha^{-1} \in C^k (1 \leqslant k \leqslant \infty)$，就称 M 是 n 维 C^k 流形。今后如不作特别说明，总假定 M 是 C^∞ 流形。

例 3 单位圆周 $C=\{(x, y) \in R^2 | x^2+y^2=1\}$，由 R^2 上的通

常拓扑诱导出C上的拓扑。考虑C的开覆盖$\{U_1, U_2\}$：

$$U_1 = \{(x, y) \in R^2 | x^2 + y^2 = 1, (x, y) \neq (1, 0)\}$$
$$= \{(x, y) \in R^2 | x = \cos\theta, y = \sin\theta, 0 < \theta < 2\pi\};$$
$$U_2 = \{(x, y) \in R^2 | x^2 + y^2 = 1, (x, y) \neq (-1, 0)\}$$
$$= \{(x, y) \in R^2 | x = \cos\theta', y = \sin\theta', -\pi < \theta' < \pi\}.$$

U_1 联系着映射 φ_1，

$$\varphi_1: U_1 \to (0, 2\pi), (x, y) = (\cos\theta, \sin\theta) \mapsto \theta,$$

U_2 联系着映射 φ_2，

$$\varphi_2: U_2 \to (-\pi, \pi), \quad (x, y) = (\cos\theta', \sin\theta') \mapsto \theta',$$

显然，φ_1 和 φ_2 都是同胚映射，所以C是1维流形。

再考察

$$\varphi_2 \circ \varphi_1^{-1}: \varphi_1(U_1 \cap U_2) \to \varphi_2(U_1 \cap U_2),$$

其中 $\varphi_1(U_1 \cap U_2) = (0, \pi) \cup (\pi, 2\pi)$，$\varphi_2(U_1 \cap U_2) = (-\pi, 0) \cup (0, \pi)$，有

$$\theta' = \varphi_2 \circ \varphi_1^{-1}(\theta) = \begin{cases} \theta, & 0 < \theta < \pi; \\ \theta - 2\pi, & \pi < \theta < 2\pi. \end{cases}$$

$\varphi_2 \circ \varphi_1^{-1}$ 和它的逆映射都可微，并且属于C^∞，所以C是1维C^∞流形。

例 4 R^3中的抛物面、椭球面、柱面、马鞍面、双曲面、车胎面等等都是2维C^∞流形。R^n中的单位球面$S = \{(x_1, \cdots, x_n) \in R^n | x_1^2 + x_2^2 + \cdots + x_n^2 = 1\}$是$n-1$维$C^\infty$流形。

要验证这些事实虽不难，但也不简单，下面的定理1给出一个较方便的判别方法。

欧几里得空间内的微分流形

在欧几里得空间中，对微分流形的陈述可以更直观和更容易掌握。设$M \subset R^n$，$\{U_\alpha\}$是R^n中的一族开集，覆盖了M，如果对每一个U_α，联系着一个同胚映射φ_α和开集$V_\alpha \subset R^n$，满足

(i) $\varphi_\alpha: U_\alpha \to V_\alpha$，又若$\varphi_\alpha$和$\varphi_\alpha^{-1}$都连续可微（这时称$\varphi_\alpha$是微分同胚），

(ii) $\varphi_\alpha: U_\alpha \cap M \to V_\alpha \cap \{(y_1, \cdots, y_n) \in R^n | y_{k+1} = y_{k+2} = \cdots = y_n = 0\}$, $x = (x_1, \cdots, x_n) \mapsto (y_1, \cdots, y_k, 0, \cdots, 0)$。

就称M是k维微分流形，$U_\alpha \cap M$是M的坐标邻域，$(y_1, \cdots, y_k) \in R^k$是点$x \in M$的局部坐标。又如果$\varphi_\alpha$和$\varphi_\alpha^{-1}$都属于$C^m$ $(1 \leqslant m \leqslant \infty)$，则称$M$是$k$维$C^m$流形。

显然，R^n中的任何一个开集必定是n维C^∞流形。

定理 1 设$f: R^n \to R$，f在R^n连续可微。令

$$M = \{x \in R^n | f(x) = 0\},$$

又设f的导数Df在M内无零点，即$1 \times n$矩阵$\left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n}\right)$在点$x \in M$的秩是1，则$M$是$n-1$维微分流形。

证明 设x_0是M内任意一点，不妨假定$\frac{\partial f}{\partial x_n}(x_0) \neq 0$，作映射

$$\varphi: R^n \to R^n, \quad x = (x_1 \cdots, x_n) \mapsto y = (y_1, \cdots, y_n),$$

$$\text{其中 } y_i = x_i,\ 1 \leqslant i \leqslant n-1,\ y_n = f(x_1, \cdots, x_n).$$

由逆映射定理，存在含有x_0的开集U和含有$y_0 = f(x_0)$的开集V，使得$\varphi: U \to V$是双射，并且存在连续可微的逆映射$\varphi^{-1}: V \to U$。

此外，当$x \in M$时，$y_n = f(x) = 0$，所以

$$\varphi: U \cap M \to V \cap \{(y_1, \cdots, y_n) | y_n = 0\},$$

这就证明了M是$n-1$维微分流形。证毕。

利用定理1容易证明例4中的抛物面、椭球面、柱面、马鞍面、双曲面等都是2维C^∞流形，R^n中的单位球面是$n-1$维C^∞流形。此外，本节后面的习题6将给出比定理1更一般的结论。

例 5 设M是所有$n \times n$矩阵组成的线性空间。对每个$A = (a_{ij}) \in M$，定义$\|A\| = \left\{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2\right\}^{1/2}$。作为赋范线性空间，$M$和$R^{n^2}$是同一个空间。令

$$GL(n, R) = \{A \in M | \det A \neq 0\},$$

则$GL(n, R)$是M中的开集，这是因为由元素$a_{11}, \cdots, a_{1n}, a_{21}, \cdots, a_{2n}, \cdots, a_{n1}, \cdots, a_{nn}$组成的行列式是变元$a_{ij}(i = 1, 2, \cdots, n;$

$j=1, 2, \cdots, n$)的连续函数，所以 $\{A\in M \mid \det M=0\}$ 是 M 中的闭集，因此 $GL(n, R)$ 是 M 中的开集，于是它是 n^2 维 C^∞ 流形。

例 6 设 l 是平面上的一条曲线，它的参数方程是

$$x=x(t)=\begin{cases}\dfrac{1}{t}, & 1\leqslant t<+\infty;\\ 0, & -\infty<t\leqslant -1.\end{cases}$$

$$y=y(t)=\begin{cases}\sin \pi t, & 1\leqslant t<+\infty;\\ t+2, & -\infty<t\leqslant -1.\end{cases}$$

当 $-1<t<1$ 时，对应的曲线是任意一条连结点(0, 1)和点(1, 0)的光滑曲线(图 4-4)。

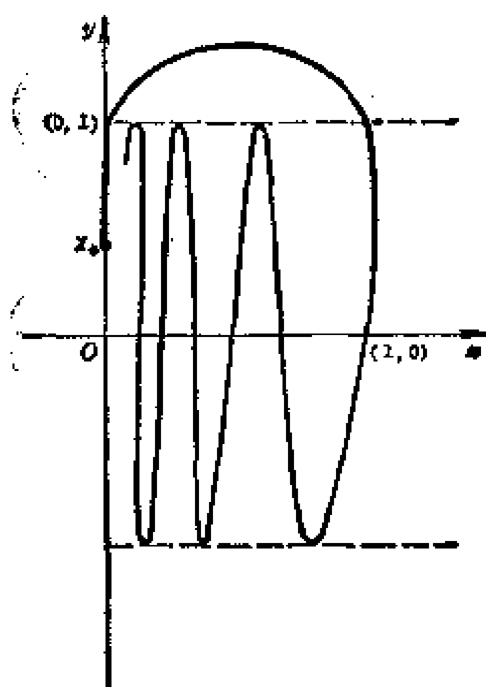


图 4-4

l 上的拓扑是由 R^2 中的通常拓扑诱导出来的，l 就成为 R^2 的一个拓扑子空间。粗看起来 l 似乎是 1 维微分流形，其实不然，它并不是 1 维流形。这是因为：如果它是 1 维流形，设 x_0 是 l 上的一点，它在 y 轴上的 -1 和 1 之间。在拓扑空间 l 内点 x_0 的邻域如图 4-5 所示，可以证明它不和直线上的任何开集同胚，这为什么？请读者考虑(作为习题)。

按照微分流形的概念再回过来看 R^3 中的光滑曲面 S，设 U

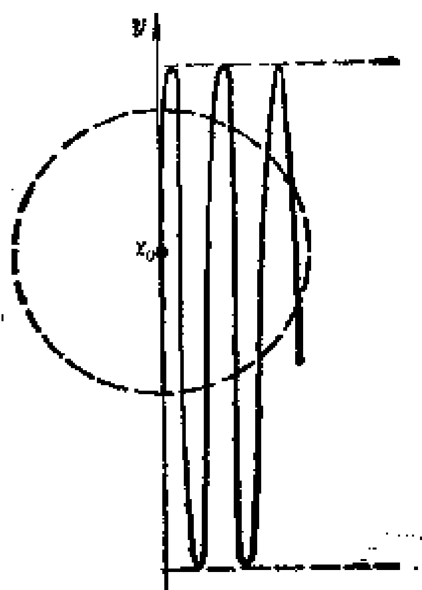


图 4-5

是 S 的一个坐标邻域，φ 是相应的坐标映射：

$$\varphi: U \to V$$

$$(x, y, z) \mapsto (u, v),$$

这就是说，在曲面 S 的一个局部范围 U 内给出了它的参数方程 φ^{-1}，即

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v),$$

其中 $(u, v) \in V$， $(x, y, z) \in U$。

微分流形上的坐标变换有下面的性质：

性质 设 M 是 n 维微分流形，$\{(U_\alpha, \varphi_\alpha)\}$ 是坐标图册，当 $U_\alpha \cap U_\beta \neq \varnothing$ 时，有

$$\det D(\varphi_\beta \circ \varphi_\alpha^{-1}) = \frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)} \neq 0,$$

$$(x_1, \cdots, x_n) \in \varphi_\alpha(U_\alpha \cap U_\beta).$$

即微分流形上的坐标邻域之间的坐标变换的雅可比行列式无零点。

证明 因为 $(\varphi_\alpha \circ \varphi_\beta^{-1}) \circ (\varphi_\beta \circ \varphi_\alpha^{-1})$ 是从 $\varphi_\alpha(U_\alpha \cap U_\beta)$ 到自身的恒等映射，所以

$$\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)} \cdot \frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)} = 1.$$

由此即得 $\det D(\varphi_\beta\circ\varphi_\alpha^{-1})\neq 0$。证毕。

定向

R^3 中的曲面可以分为两类：一类是双侧曲面；另一类是单侧曲面。例如球面、抛物面、车胎面等都是双侧曲面；牟比乌斯(Möbius)带是单侧曲面。现在将这一概念拓广到流形上去。

设 M 是 n 维微分流形，$\{(U_\alpha, \varphi_\alpha)\}$ 是它的坐标图册，对于 $U_\alpha\cap U_\beta\neq\emptyset$ 的每一对 α、β，坐标变换

$$\varphi_\beta\circ\varphi_\alpha^{-1}:\varphi_\alpha(U_\alpha\cap U_\beta)\to\varphi_\beta(U_\alpha\cap U_\beta)$$
$$x=(x_1, \cdots, x_n)\mapsto y=(y_1, \cdots, y_n)$$

的雅可比行列式如果恒正，即

$$\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}>0, \quad (x_1, \cdots, x_n)\in U_\alpha\cap U_\beta,$$

就称 M 是定向的微分流形。

例如 R^3 中的二维光滑曲面 S 可以用 U_α 和 U_β 覆盖，在坐标图 $(U_\alpha, \varphi_\alpha)$ 内，曲面 S 表示为

$$x=x(u, v), \quad y=y(u, v), \quad z=z(u, v),$$

其中 $(x, y, z)\in U_\alpha$，$(u, v)\in\varphi_\alpha(U_\alpha)$，它的单位法向量是

$$n=\frac{r_u\times r_u}{|r_u\times r_v|},$$
$$r_u=(x_u, y_u, z_u),$$
$$r_v=(x_v, y_v, z_v).$$

在坐标图 (U_β, φ_β) 内，曲面 S 表示为

$$x=x(s, t), \quad y=y(s, t), \quad z=z(s, t),$$

其中 $(x, y, z)\in U_\beta$，$(s, t)\in\varphi_\beta(U_\beta)$，它的单位法向量是

$$n=\frac{r_s\times r_t}{|r_s\times r_t|},$$
$$r_s=(x_s, y_s, z_s),$$
$$r_t=(x_t, y_t, z_t).$$

由 $\varphi_\beta\circ\varphi_\alpha^{-1}$ 建立了从 (u, v) 到 (s, t) 的坐标变换：

$$s=s(u, v), \quad t=t(u, v),$$

变换的雅可比行列式是$\frac{\partial(s,t)}{\partial(u,v)}$，在这一变换下，$r_u\times r_v$ 和$r_s\times r_t$ 有下面的关系：

$$r_u\times r_v=\frac{\partial(y,z)}{\partial(u,v)}i+\frac{\partial(z,x)}{\partial(u,v)}j+\frac{\partial(x,y)}{\partial(u,v)}k$$

$$=\frac{\partial(s,t)}{\partial(u,v)}\left(\frac{\partial(y,z)}{\partial(s,t)}i+\frac{\partial(z,x)}{\partial(s,t)}j\right.$$

$$\left.+\frac{\partial(x,y)}{\partial(s,t)}k\right)=\frac{\partial(s,t)}{\partial(u,v)}r_s\times r_t.$$

如果$\frac{\partial(s,t)}{\partial(u,v)}>0$，即曲面$S$是定向的，这表明$U_\alpha$中的法方向和$U_\beta$中的法方向是一致的，或者说在定向的光滑曲面上，当法线沿闭曲面上的连续闭曲线变动一周后回到原来的位置时，法线的方向和原来的方向是一致的。

例 7 设$S=\{(x,y,z)\in R^3|x^2+y^2+z^2=1\}$，考察$S$的六个坐标邻域

$U_1^+=\{(x,y,z)\in S|x>0\}$， 坐标是 y，z；
$U_1^-=\{(x,y,z)\in S|x<0\}$， 坐标是 z，y；
$U_2^+=\{(x,y,z)\in S|y>0\}$， 坐标是 z，x；
$U_2^-=\{(x,y,z)\in S|y<0\}$， 坐标是 x，z；
$U_3^+=\{(x,y,z)\in S|z>0\}$， 坐标是 x，y；
$U_3^-=\{(x,y,z)\in S|z<0\}$， 坐标是 y，x。

(坐标系的选取服从右手法则)，在$U_1^+\cap U_2^+$内，坐标变换是

$$\begin{cases}y=\sqrt{1-z^2-x^2};\\ z=z,\quad x>0,\quad y>0,\end{cases}$$

$$\frac{\partial(y,z)}{\partial(z,x)}=\begin{vmatrix}\frac{-z}{\sqrt{1-z^2-x^2}} & \frac{-x}{\sqrt{1-z^2-x^2}}\\ 1 & 0\end{vmatrix}=\frac{x}{\sqrt{1-z^2-x^2}}>0.$$

在$U_1^+\cap U_3^-$内，坐标变换为

$$\begin{cases}y=y;\\ z=-\sqrt{1-y^2-x^2},\quad x>0,\quad z<0,\end{cases}$$

$$\frac{\partial(y, z)}{\partial(y, x)}=\begin{vmatrix} 1 & 0 \\ \frac{y}{\sqrt{1-y^2-x^2}} & \frac{x}{\sqrt{1-y^2-x^2}} \end{vmatrix}$$

$$=\frac{x}{\sqrt{1-y^2-x^2}}>0.$$

在 $U_2^- \cap U_3^-$ 内，坐标变换为

$$\begin{cases} x=x; \\ z=-\sqrt{1-y^2-x^2}, \quad y<0, \quad z<0. \end{cases}$$

$$\frac{\partial(x, z)}{\partial(y, x)}=\begin{vmatrix} 0 & 1 \\ \frac{y}{\sqrt{1-y^2-x^2}} & \frac{x}{\sqrt{1-y^2-x^2}} \end{vmatrix}$$

$$=\frac{-y}{\sqrt{1-y^2-x^2}}>0.$$

在 $U_1^+ \cap U_2^-$，$U_1^+ \cap U_3^+$，… 内可作同样的讨论。由此可见 S 是定向的微分流形。

流形上的可微函数

设 M 是 n 维 C^∞ 流形，坐标图册是 $\{(U_\alpha, \varphi_\alpha)\}$，$G$ 是 M 中的一个开集，它不一定是坐标邻域，$p\in G$。又设 $f: G\to R$，何谓 f 在点 p 可微？何谓 f 在 G 可微？

考察 $f\circ\varphi_\alpha^{-1}$，它是定义在 $\varphi_\alpha(G\cap U_\alpha)$ 上的实值函数，而 $\varphi_\alpha(G\cap U_\alpha)$ 是 R^n 内的一个开集，因此 $f\circ\varphi_\alpha^{-1}$ 是 n 元实值函数。设 $x=\varphi_\alpha(p)$，即 $x=(x_1, \cdots, x_n)$ 是点 p 的局部坐标，如果 $f\circ\varphi_\alpha^{-1}$ 在点 x 可微，就称 f 在点 p 可微。如果 f 在每一点 $p\in G$ 可微，称 f 在 G 可微(图 4-6)。

自然会产生一个问题：如果 $p\in G\cap U_\alpha$ 同时 $p\in G\cap U_\beta$，利用坐标图 $(U_\alpha, \varphi_\alpha)$ 定义了 f 在 p 可微即 $f\circ\varphi_\alpha^{-1}$ 在 x 可微，那么在坐标图 (U_β, φ_β) 内，设 $y=\varphi_\beta(p)$，$f\circ\varphi_\beta^{-1}$ 是否也在点 y 可微？换句话说，可微性是否和局部坐标的选取无关？由

$$f\circ\varphi_\beta^{-1}=(f\circ\varphi_\alpha^{-1})\circ(\varphi_\alpha\circ\varphi_\beta^{-1})$$

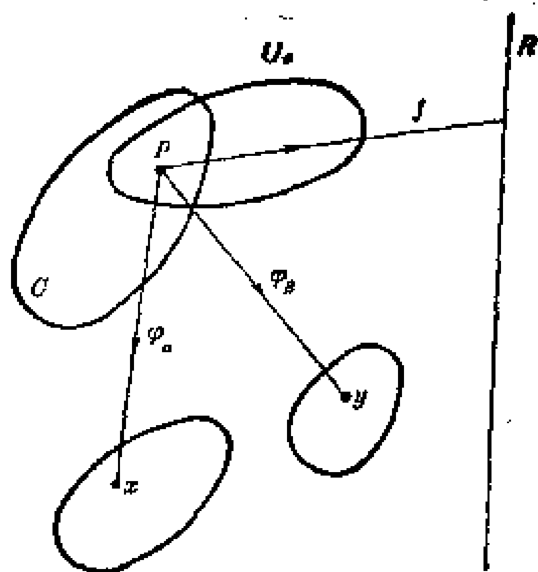


图 4-6

知道：当 $f\circ\varphi_\alpha^{-1}$ 在点 x 可微时，$f\circ\varphi_\beta^{-1}$ 也在点 y 可微。所以可微性与局部坐标的选取无关。

在上面的定义中，如果 f 在 G 可微，并且每一个 $f\circ\varphi_\alpha^{-1}\in C^\infty$，就称 f 是 C^∞ 的，记为 $f\in C^\infty$（当然这要求 M 是 C^∞ 流形）。同样可以定义 $f\in C^k$，即 M 是 C^k 流形，同时每一个 $f\circ\varphi_\alpha^{-1}\in C^k$。

例 6 设 S 是 R^3 中的单位球面，在 S 上定义了一个实值函数 f，$p=(x_0, y_0, z_0)$ 是 S 上的一点，U 是含点 p 的坐标邻域，坐标映射是 φ，再设

$$\varphi^{-1}:(u,\ v)\mapsto(x,\ y,\ z)$$

$$x=\sin u\cos v,\quad y=\sin u\sin v,$$

$$z=\cos u\quad(0<u<\pi,\ 0<v<2\pi),$$

又设 $(u_0, v_0)=\varphi(x_0, y_0, z_0)$，则

$$f\circ\varphi^{-1}(u,\ v)=f(\sin u\cos v,\ \sin u\sin v,\ \cos u),$$

由流形上函数的可微概念可知道，如果 f 作为 u、v 的函数在点 (u_0, v_0) 可微，就说 f 作为 S 上的函数在点 $p=(x_0, y_0, z_0)$ 可微。

切向量和切空间

怎样把 R^3 中光滑曲线的切线和光滑曲面的切平面的概念拓

广到微分流形上？先分析一个最简单的情形——平面上光滑曲线的切线。求切线是和“求导”紧密联系的，设 l 是平面上的一条光滑曲线，它的方程是 $y=f(x)$。$p=(x_0, y_0)$ 是 l 上的一点，求曲线 l 在点 p 的切线的关键就是求函数 f 在点 x_0 的导数 $\frac{d}{dx}f\Big|_{x_0}$，现在我们把 $\frac{d}{dx}\Big|_{x_0}$ 看作一个映射，它作用在函数 f 上就获得一个实数 $\frac{d}{dx}f\Big|_{x_0}$，这个实数就是曲线 $y=f(x)$ 在点 p 的切线斜率，它表征着切线的方向。

用数学的语言写下来就是：设 F_{x_0} 是所有在点 x_0 可微的函数组成的空间，带有通常的函数的“加”、“数乘”和“乘”的运算，即

$$(f+g)(x)=f(x)+g(x),$$

$$(\alpha f)(x)=\alpha f(x),$$

$$(fg)(x)=f(x)g(x),$$

作映射（它是在点 x_0 求导）

$$\frac{d}{dx}\Big|_{x_0}: F_{x_0}\to R$$

$$f\mapsto\frac{d}{dx}f\Big|_{x_0},$$

由求导法则知道，它满足以下两个性质：

(i) 线性：

$$\frac{d}{dx}(f+g)\Big|_{x_0}=\frac{d}{dx}f\Big|_{x_0}+\frac{d}{dx}g\Big|_{x_0}$$

$$\frac{d}{dx}(\alpha f)\Big|_{x_0}=\alpha\frac{d}{dx}f\Big|_{x_0}\quad(\alpha\text{ 是实数}),$$

(ii) 莱布尼茨(Leibniz)法则：

$$\frac{d}{dx}(fg)\Big|_{x_0}=f(x_0)\frac{d}{dx}g\Big|_{x_0}+g(x_0)\frac{d}{dx}f\Big|_{x_0}.$$

反过来，如果有一个映射

$$\varphi: F_{x_0}\to R$$

满足

(i) 线性： $\varphi(f+g)=\varphi(f)+\varphi(g)$，

$$\varphi(\alpha f)=\alpha\varphi(f).$$

(ii) 莱布尼茨法则

$$\varphi(fg)=f(x_0)\varphi(g)+g(x_0)\varphi(f).$$

那么这个 φ 是否一定是求导？

对任何 $f\in F_{x_0}$，将它展开为

$$f(x)=f(x_0)+(x-x_0)g(x),$$

其中

$$g(x)=\begin{cases}\dfrac{f(x)-f(x_0)}{x-x_0}, & x\neq x_0;\\ \dfrac{d}{dx}f\Big|_{x_0}, & x=x_0,\end{cases}$$

由 φ 的线性和莱布尼茨法则

$$\begin{aligned}\varphi(f)&=\varphi(f(x_0))+(x_0-x_0)\varphi(g)+\varphi(x-x_0)g(x_0)\\&=\varphi(f(x_0))+\varphi(x)g(x_0)-\varphi(x_0)g(x_0),\end{aligned}$$

再由莱布尼茨法则知道，对常数函数 1，

$$\varphi(1)=\varphi(1\cdot 1)=\varphi(1)+\varphi(1)=2\varphi(1),$$

所以 $\varphi(1)=0$，于是对任何常数函数 c，

$$\varphi(c)=c\varphi(1)=0,$$

代入 $\varphi(f)$，有

$$\varphi(f)=\varphi(x)g(x_0)=\varphi(x)\frac{d}{dx}f\Big|_{x_0},$$

$\varphi(x)$ 是一个常数，它与 f 无关，只与 φ 有关，记它是 α，所以

$$\varphi(f)=\alpha\frac{d}{dx}f\Big|_{x_0},$$

这表明映射 φ 作用在任何 $f\in F_{x_0}$ 上，等于求导 $\alpha\dfrac{d}{dx}\Big|_{x_0}$ 作用在 f 上，即

$$\varphi\equiv\alpha\frac{d}{dx}\Big|_{x_0}.$$

系数 α 是一个与 f 无关只与 φ 有关的常数。这又说明所有这种 φ

组成一个1维的线性空间，基是$\left.\frac{d}{dx}\right|_{x_0}$，每一个这样的$\varphi$作用在$f\in F_{x_0}$上，便得到曲线$y=f(x)$在点$(x_0, f(x_0))$的切线方向，称$\varphi$是切向量。

切向量φ是一个满足线性和莱布尼茨法则的映射，读者也许会觉得不太自然，因为我们习惯于把向量看成是有方向和大小的量，而不是映射．其实，把它看作映射或许更自然和有力。例如在童话故事中人们把金银珠宝看作财富，但“点金术”也是财富，映射φ就是“点金术”，把它作用在任何可微函数上就得到一个切向量，那么把φ称为切向量不是也很自然吗？这好比你愿意背一大框金银珠宝呢？还是愿意掌握“点金术”呢？



现在，我们将这一基本思想运用到微分流形上．设M是n维C^∞流形，$p\in M$．又设F_p是所有在点p附近有定义并且在p可微的函数组成的空间，带有下列的“加”、“数乘”和“乘”的运算：

$$(f+g)(p')=f(p')+g(p'),$$
$$(\alpha f)(p')=\alpha f(p'),$$
$$(fg)(p')=f(p')\cdot g(p').$$

设映射

$$\varphi: F_p\to R$$

满足(i)线性：
$$\varphi(f+g)=\varphi(f)+\varphi(g),$$
$$\varphi(\alpha f)=\alpha\varphi(f).$$

(ii) 莱布尼茨法则：

$$\varphi(fg)=f(p)\varphi(g)+g(p)\varphi(f).$$

就称φ是M在点p的一个切向量，所有这种φ组成的空间称为M在点p的切空间，记为T_p，其中的“加”和“数乘”是：对任何$f\in F_p$，任何φ_1、$\varphi_2\in T_p$，以及任何实数α，成立

$$(\varphi_1+\varphi_2)(f)=\varphi_1(f)+\varphi_2(f),$$
$$\varphi_1(\alpha f)=\alpha\varphi_1(f).$$

装备了上述“加”和“数乘”之后，T_p是一个线性空间．那么φ具有怎样的形式？T_p的基是什么？

定理 2 切空间 T_p 是 n 维线性空间，在给定的坐标系 $(x_1, \cdots, x_n)$ 下，它的基是

$$\left.\frac{\partial}{\partial x_1}\right|_{x_0}, \cdots, \left.\frac{\partial}{\partial x_n}\right|_{x_0},$$

其中 $x_0=(x_1^0, \cdots, x_n^0)$，它是点 p 的局部坐标。每一个切向量 φ 都可以用 $\left.\frac{\partial}{\partial x_j}\right|_{x_0}$ $(j=1, 2, \cdots, n)$ 线性表示出：

$$\varphi=\alpha_1 \left.\frac{\partial}{\partial x_1}\right|_{x_0}+\cdots+\alpha_n \left.\frac{\partial}{\partial x_n}\right|_{x_0}.$$

证明 设 $(U_\alpha, \varphi_\alpha)$ 是坐标图，$p \in U_\alpha$。对任何 $f \in F_p$，f 在点 p 可微，即 $f \circ \varphi_\alpha^{-1}$ 在点 x_0 可微，为了使记号简单起见，我们把 $f \circ \varphi_\alpha^{-1}$ 仍旧记为 f。将 f 展开：

$$f(x)=f(x_0)+\sum_{i=1}^{n}(x_i-x^0)g_i(x),$$

其中 g_i 在点 x_0 的值是

$$g_i(x_0)=\left.\frac{\partial}{\partial x_i} f\right|_{x_0}.$$

切向量 φ 作用于 f，得

$$\begin{aligned}\varphi(f)&=\varphi(f(x_0))+\sum_{i=1}^{n}(x^0-x_i^0)\varphi(g_i)+\sum_{i=1}^{n}\varphi(x_i-x_i^0)g_i(x_0)\\&=\varphi(f(x_0))+\sum_{i=1}^{n}\varphi(x_i)g_i(x_0)-\sum_{i=1}^{n}\varphi(x_i^0)g_i(x_0),\end{aligned}$$

由莱布尼茨法则：

$$\varphi(1)=\varphi(1\cdot 1)=\varphi(1)+\varphi(1),$$

所以 $\varphi(1)=0$，从而对任何常量函数 c，$\varphi(c)=0$。

$$\varphi(f)=\sum_{i=1}^{n}\varphi(x_i)g_i(x_0)=\sum_{i=1}^{n}\varphi(x_i)\left.\frac{\partial}{\partial x_i} f\right|_{x_0},$$

记 $\alpha_i=\varphi(x_i)$ $(i=1, 2, \cdots, n)$，α_i 与 f 无关，它们是由 φ 确定的，则对任何 $f \in F_p$，

$$\varphi(f)=\sum_{i=1}^{n}\alpha_i \left.\frac{\partial}{\partial x_i} f\right|_{x_0},$$

即

$$\varphi=\sum_{i=1}^{n}\alpha_i\left.\frac{\partial}{\partial x_i}\right|_{x_0},$$

这表明切向量 φ 是 $\left.\frac{\partial}{\partial x_i}\right|_{x_0}$ $(i=1, 2, \cdots, n)$ 的线性组合。

而每一个 $\left.\frac{\partial}{\partial x_i}\right|_{x_0}$ 是满足线性和莱布尼茨法则的，因此 $\left.\frac{\partial}{\partial x_i}\right|_{x_0}$ 是一个切向量。

再证明 $\left.\frac{\partial}{\partial x_1}\right|_{x_0}, \cdots, \left.\frac{\partial}{\partial x_n}\right|_{x_0}$ 线性无关。设存在一组实数 $c_1, \cdots, c_n$，使

$$c_1\left.\frac{\partial}{\partial x_1}\right|_{x_0}+\cdots+c_n\left.\frac{\partial}{\partial x_n}\right|_{x_0}=0,$$

(右端的 0 是 0 映射)，将它作用在函数 x_i 上：

$$\left(c_1\left.\frac{\partial}{\partial x_1}\right|_{x_0}+\cdots+c_n\left.\frac{\partial}{\partial x_n}\right|_{x_0}\right)(x_i)=0,$$

得 $c_i=0$, $(i=1, 2, \cdots, n)$，所以 $\left.\frac{\partial}{\partial x_1}\right|_{x_0}, \cdots, \left.\frac{\partial}{\partial x_n}\right|_{x_0}$ 线性无关，这样便获得定理的全部结论。证毕。

习　题

1. 证明开区间是 1 维流形，闭区间不是 1 维流形。

2. 证明上半锥面 $z^2=x^2+y^2$, $z\geqslant 0$ 是 2 维流形，但整个锥面不是 2 维流形。

3. 证明 R^3 中的抛物面、马鞍面、椭球面、双曲面、柱面都是 2 维 C^∞ 流形。

4. 证明例 6 中的曲线 l 不是 1 维流形。

5. 设 $SL(n, R)$ 是所有行列式为 1 的 $n\times n$ 矩阵组成的集合，对 $A=(a_{ij})\in SL(n, R)$，定义 $\|A\|=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2\right)^{1/2}$。证明 $SL(n, R)$ 是 n^2-1 维 C^∞ 流形。

6. 设 f 是从 R^n 到 R^k 的连续可微映射，$M=\{x\in R^n | f(x)=0\}$，又设 Df 在点 $x\in M$ 的秩是 k, $k<n$。证明 M 是 $n-k$ 维微分流形。

7．证明 $M=\{(x, y, u, v)\in R^4 \mid x^2+y^2=1,\ u^2+v^2=1\}$ 是2维微分流形。

8．验证平面上的单位圆周 $C=\{(x, y)\mid x^2+y^2=1\}$ 是定向的流形，并说明其几何意义。

第 5 章　微分形式和外微分

§5.1　外　积

外积是一种代数运算，它具有相当直观的几何背景。本节的中心是从给定的线性空间出发，构造出一个新的线性空间，使后者不仅具有线性空间的代数结构，并且还具有一个新的代数运算——外积，为即将引进的微分形式作准备。

直观的几何背景

这不是严格的论述，而是借助于几何的直观，为我们提供一个背景材料，启示我们如何去思考和去发现。

在 3 维欧几里得空间 R^3 内，设向量

$$a_1 = a_{11}i + a_{12}j + a_{13}k,$$
$$a_2 = a_{21}i + a_{22}j + a_{23}k,$$
$$a_3 = a_{31}i + a_{32}j + a_{33}k,$$

并假设 a_1, a_2, a_3 线性无关，那么

$$V = \{x \in R^3 | x = \alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3,\ 0 \leqslant \alpha_1,\ \alpha_2,\ \alpha_3 \leqslant 1\}$$

就是由向量 a_1, a_2, a_3 张成的平行 6 面体（图 5-1）。在 a_1, a_2, a_3 之间引进一个运算$\wedge$如下：

$$a_1 \wedge a_2 \wedge a_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

由解析几何知道：$a_1 \wedge a_2 \wedge a_3$ 的几何意义是 V 的有向容积。所谓"有向"是指这个容积是带有正负号的，在给定的服从右手法则的笛卡儿直角坐标系内(这就是通常的坐标系)，当 a_1, a_2, a_3 的顺序服从右手法则时，上述容积带正号；否则，当 a_1, a_2, a_3 的顺序

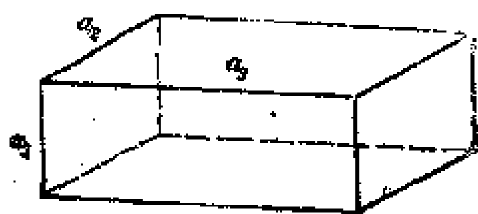


图 5-1

服从左手法则时，容积带负号。

容易看出运算$\wedge$服从下列法则：

(i) 重线性：在 $a_1\wedge a_2\wedge a_3$ 中，如果某一个向量(例如 a_2)是其他向量的线性组合(例如 $a_2=\beta b+\gamma c$，b、c 是向量，β、γ 是实数)，那么

$$a_1\wedge(\beta b+\gamma c)\wedge a_3=\beta(a_1\wedge b\wedge a_3)+\gamma(a_1\wedge c\wedge a_3).$$

(ii) 反交换性：在 $a_1\wedge a_2\wedge a_3$ 中，任何两个向量交换之后符号相反，即

$$a_1\wedge a_2\wedge a_3=-a_2\wedge a_1\wedge a_3,$$
$$a_1\wedge a_2\wedge a_3=-a_3\wedge a_2\wedge a_1,$$
$$a_1\wedge a_2\wedge a_3=-a_1\wedge a_3\wedge a_2,$$

由 (ii) 又知道，在 $a_1\wedge a_2\wedge a_3$ 中如果有两个向量相同，那么

$$a_1\wedge a_2\wedge a_3=0.$$

例 1 设 e_1，e_2，e_3 是 R^3 内一组基，它们不一定正交，又设 a_1，a_2，a_3 是 R^3 内的三个向量，在基 e_1，e_2，e_3 下表示为

$$a_1=a_{11}e_1+a_{12}e_2+a_{13}e_3,$$
$$a_2=a_{21}e_1+a_{22}e_2+a_{23}e_3,$$
$$a_3=a_{31}e_1+a_{32}e_2+a_{33}e_3,$$

由运算$\wedge$的重线性和反交换性，经过直接计算，立即得到

$$a_1\wedge a_2\wedge a_3=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}e_1\wedge e_2\wedge e_3.$$

对这一等式作一番几何的解释是很值得的。上式的左端表示由 a_1，a_2，a_3 所张平行 6 面体的有向容积；右端的 e_1，e_2，e_3 是由基向量 e_1，e_2，e_3 所张平行 6 面体的有向容积，行列式是这两个容积之间带有符号的比例系数，行列式符号为正，即 a_1，a_2，a_3 的顺序和 e_1，e_2，e_3 的顺序是一致的，例如都服从右手法则，或都服从左手法则；行列式符号为负，即 a_1，a_2，a_3 的顺序和 e_1，e_2，e_3 的顺序相反，其中一个服从右手法则，另一个服从左手法则。

例 2　设 e_1, e_2, e_3 是 R^3 内的一组基. 又设

$$a_1 = a_{11}e_1 + a_{12}e_2 + a_{13}e_3,$$

$$a_2 = a_{21}e_1 + a_{22}e_2 + a_{23}e_3,$$

在 a_1 和 a_2 之间引进运算 $\wedge$，要求它满足重线性和反交换性，那么

$$a_1 \wedge a_2 = (a_{11}e_1 + a_{12}e_2 + a_{13}e_3) \wedge (a_{21}e_1 + a_{22}e_2 + a_{23}e_3)$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} e_1 \wedge e_2 + \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} e_2 \wedge e_3$$

$$+ \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix} e_3 \wedge e_1.$$

对这一等式作几何解释也是很有意义的。在 R^3 内我们把 e_1, e_2, e_3 看作坐标向量，并由它们产生坐标平面，例如 e_1e_2–面就是由坐标向量 e_1、e_2 产生的平面（即包含 e_1 和 e_2 的平面）。记 a_1、a_2 张成的平行四边形是 A，即

$$A = \{x \in R^3 \mid x = \alpha_1 a_1 + \alpha_2 a_2,\ 0 \leqslant \alpha_1,\ \alpha_2 \leqslant 1\},$$

A 在 e_1e_2–面上的投影记为 A_{12}，A 在 e_2e_3–面和 e_3e_1–面上的投影分别记为 A_{23} 和 A_{31}，这些投影都是平行四边形或退化为一条直线段，等式右端各项表示 A 在各坐标平面上的投影面积，例如第一项是投影 A_{12} 的有向面积，它是用带符号的比例系数 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ 乘以 e_1、e_2 生成的平行四边形的有向面积来表示的。比例系数的符号是这样确定的：如我们取坐标 e_1、e_2 的顺序是正向［如同我们在笛卡儿坐标系内取 x、y 的顺序（x 在前，y 在后）是正向一样］，那么当向量 a_1、a_2 在 e_1e_2–面上的投影为 a_1' 和 a_2'（它们都是向量），在由 a_1'、a_2' 所张的平行四边形内部从 a_1' 旋转到 a_2' 时，其旋转方向和 e_1、e_2 的顺序一致，那么比例系数的符号是正，这时投影 A_{12} 的符号也是正的（图 5-2）；否则，符号为负。

将运算 $\wedge$ 所服从的重线性和反交换性作为出发点，加以抽象，便得到外积的概念。

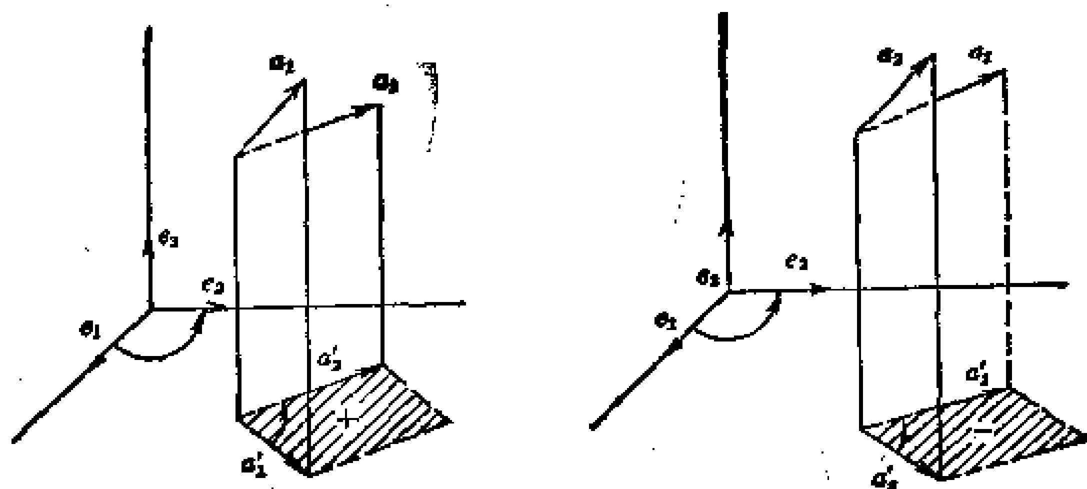


图 5-2

外积·格拉斯曼代数

设R是实数域，V 是 n 维实线性空间，V 中的元素用 u,v 表示，基是 $e_1, e_2, \cdots, e_n$，对 $k=0, 1, 2, \cdots, n$，我们构造实数域上新的线性空间 Λ^k 如下：

当 $k=0$、1时，定义

$$\Lambda^0 = R, \qquad \Lambda^1 = V,$$

定义 Λ^2 是由所有形如

$$\sum a_i(u_i \wedge v_i)$$

的有限项的和式组成，其中 $\alpha_i \in R$，u_i、$v_i \in V$，并要求$\wedge$满足下列条件(其中 α_1、$\alpha_2 \in R$，u_1、u_2、v_1、v_2、u、$v \in V$)：

(i) 重线性：

$$(\alpha_1 u_1 + \alpha_2 u_2) \wedge v = \alpha_1(u_1 \wedge v) + \alpha_2(u_2 \wedge v),$$
$$u \wedge (\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1(u \wedge v_1) + \alpha_2(u \wedge v_2).$$

(ii) 反交换性：

$$u \wedge v = -v \wedge u.$$

(iii) $u \wedge u = 0$.

$u \wedge v$ 称为是 u 和 v 的外积(或楔积)。如果 u、v 非独立，设 $u = av$，则

$$u\wedge v=\alpha(v\wedge v)=0.$$

在给定的一组基 $e_1, e_2, \cdots, e_n$ 之下，设

$$u=\sum_{i=1}^{n}\alpha_i e_i,\ v=\sum_{i=1}^{n}\beta_i e_i,$$

则

$$u\wedge v=\left(\sum_{i=1}^{n}\alpha_i e_i\right)\wedge\left(\sum_{i=1}^{n}\beta_i e_i\right)$$

$$=\sum_{i,j}\alpha_i\beta_j(e_i\wedge e_j)$$

注意到 $e_i\wedge e_i=0$, $e_j\wedge e_i=-e_i\wedge e_j$，于是，

$$u\wedge v=\sum_{i<j}(\alpha_i\beta_j-\alpha_j\beta_i)e_i\wedge e_j.$$

由此可见 Λ^2 中元素的标准形式是下列外积的线性组合：

$$e_i\wedge e_j\quad(1\leqslant i<j\leqslant n).$$

Λ^2 的元素称为是 2-形式（并分别称 Λ^0 和 Λ^1 的元素是 0-形式和 1-形式），这就是说，线性空间 Λ^2 的基是 $e_i\wedge e_j$ $(1\leqslant i<j\leqslant n)$，$\Lambda^2$ 的维数是 $\binom{n}{2}=\frac{n(n-1)}{2}$，其元素的一般形式是

$$\sum_{1\leqslant i<j\leqslant n}\alpha_{ij}e_i\wedge e_j\quad(\alpha_{ij}\in R).$$

一般的，可以相仿给出 Λ^k $(2<k\leqslant n)$，它是由所有形如

$$\sum\alpha_{i_1\cdots i_k}(v_{i_1}\wedge v_{i_2}\wedge\cdots\wedge v_{i_k})$$

的有限项和式所组成，其中 $\alpha_{i_1\cdots i_k}\in R$, $v_{i_1},\cdots,v_{i_k}\in V$. $\wedge$ 满足下列条件：

(i) 重线性：

$$v_{i_1}\wedge\cdots\wedge(\alpha v_{i_j}+\alpha' v'_{i_j})\wedge\cdots\wedge v_{i_k}$$

$$=\alpha v_{i_1}\wedge\cdots\wedge v_{i_j}\wedge\cdots\wedge v_{i_k}+\alpha' v_{i_1}\wedge\cdots\wedge v'_{i_j}\wedge\cdots\wedge v_{i_k},$$

其中 $\alpha, \alpha'\in R$; $v_{i_1},\cdots,v_{i_j}, v'_{i_j},\cdots,v_{i_k}\in V$.

(ii) 反交换性：

$$v_{i_1}\wedge\cdots\wedge v_{i_j}\wedge\cdots\wedge v_{i_l}\wedge\cdots\wedge v_{i_k}$$

$$=-v_{i_1}\wedge\cdots\wedge v_{i_l}\wedge\cdots\wedge v_{i_j}\wedge\cdots\wedge v_{i_k}.$$

(iii) 如果在 $v_{i_1},\cdots,v_{i_k}$ 中有两个元素相同，则

$$v_{i_1}\wedge\cdots\wedge v_{i_p}=0.$$

在给定的基 $e_1, e_2, \cdots, e_n$ 下，V 中 k 个元素 $v_{i_1}, \cdots, v_{i_k}$ 的外积是

$$v_{i_1}\wedge\cdots\wedge v_{i_k}=\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\beta_{j_1\cdots j_k}e_{j_1}\wedge\cdots\wedge e_{j_k},$$

其中
$$\beta_{j_1\cdots j_k}\in R.$$

可见 Λ^k 的基是

$$e_{i_1}\wedge\cdots\wedge e_{i_k}\quad(1\leqslant i_1<\cdots<i_k\leqslant n),$$

Λ^k的维数是$\binom{n}{k}$，Λ^k 中元素的一般形式是

$$\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}a_{i_1\cdots i_k}e_{i_1}\wedge\cdots\wedge e_{i_k},$$

$a_{i_1\cdots i_k}$ 是实数。这一形式称为 *k-形式*。

当 $p>n$ 时，$v_{i_1}\wedge\cdots\wedge v_{i_p}$ 必为 0，这是因为 $v_{i_j}(1\leqslant j\leqslant p)$ 用基 $e_1, e_2, \cdots, e_n$ 线性表示之后，将外积 $v_{i_1}\wedge\cdots\wedge v_{i_p}$ 展开，展开式的每一项 $e_{i_1}\wedge\cdots\wedge e_{i_p}(p>n)$ 中至少有两个元素是相同的，所以等于 0。

当 $p=n$ 时，Λ^n 是 1 维线性空间，基是 $e_1\wedge e_2\wedge\cdots\wedge e_n$，其元素的标准形式是

$$\alpha e_1\wedge e_2\wedge\cdots\wedge e_n\quad(\alpha\in R).$$

到这里为止，我们从线性空间 V 出发，定义了一列新的线性空间 $\Lambda^k(1\leqslant k\leqslant n)$，并在每个 Λ^k 内定义了外积。现在再在 Λ^p 和 Λ^q 之间($p\geqslant1$，$q\geqslant1$；$p+q\leqslant n$)定义外积 $\wedge$，设 ω 是一个 p-形式，η 是一个 q-形式，定义 $\omega\wedge\eta$ 是$(p+q)$-形式，即

$$\wedge: \Lambda^p\times\Lambda^q\to\Lambda^{p+q}\quad(p+q\leqslant n),$$
$$(\omega, \eta)\mapsto\omega\wedge\eta.$$

$\omega\wedge\eta$ 是将 ω 和 η 各自用基表示后再作$\wedge$运算。

对任何实数 α(即 $\alpha\in\Lambda^0$)和任何形式 ω，定义 $\alpha\wedge\omega=\omega\wedge\alpha=\alpha\omega$，即数 α 乘 ω。

例 3 设

$$\omega=Ae_1+Be_2+Ce_3\quad(1\text{-形式}),$$
$$\eta=Pe_2\wedge e_3+Qe_3\wedge e_1+Re_1\wedge e_2\quad(2\text{-形式}).$$

经过直接计算得

$$\omega\wedge\eta=(AP+BQ+CR)e_1\wedge e_2\wedge e_3.$$

外积又具有以下性质：

性质 1 设 ω 是 p-形式，η 是 q-形式，$p+q\leqslant n$，那么

$$\omega\wedge\eta=(-1)^{pq}\eta\wedge\omega.$$

证明 由外积的重线性，只要证明

$$\omega=e_{i_1}\wedge\cdots\wedge e_{i_p};$$

$$\eta=e_{j_1}\wedge\cdots\wedge e_{j_q}$$

的情形就可以了．这时

$$\begin{aligned}\omega\wedge\eta&=(e_{i_1}\wedge\cdots\wedge e_{i_p})\wedge(e_{j_1}\wedge\cdots\wedge e_{j_q})\\&=(-1)^p e_{j_1}\wedge e_{i_1}\wedge\cdots\wedge e_{i_p}\wedge e_{j_2}\wedge\cdots\wedge e_{j_q}\\&=(-1)^{2p}e_{j_1}\wedge e_{j_2}\wedge e_{i_1}\wedge\cdots\wedge e_{i_p}\wedge e_{j_3}\wedge\cdots\wedge e_{j_q}\\&=\cdots\cdots\\&=(-1)^{pq}e_{j_1}\wedge\cdots\wedge e_{j_q}\wedge e_{i_1}\wedge\cdots\wedge e_{i_p}.\end{aligned}$$

证毕

性质 2 设

$$v_i=a_{i1}e_1+\cdots+a_{in}e_n\quad(i=1,\ 2,\ \cdots,\ m;\ m\leqslant n),$$

那么

$$v_1\wedge\cdots\wedge v_m=\sum_{1\leqslant i_1<\cdots<i_m\leqslant n}\begin{vmatrix}a_{1i_1}&\cdots&a_{1i_m}\\\cdots&\cdots&\cdots\\a_{mi_1}&\cdots&a_{mi_m}\end{vmatrix}e_{i_1}\wedge\cdots\wedge e_{i_m}.$$

特别是，若 $m=n$，那么

$$v_1\wedge\cdots\wedge v_n=\begin{vmatrix}a_{11}&\cdots&a_{1n}\\\cdots&\cdots&\cdots\\a_{n1}&\cdots&a_{nn}\end{vmatrix}e_1\wedge\cdots\wedge e_n.$$

证明 为书写方便起见，不失一般性，设 $n=3$，$m=2$，这时

$$v_1=a_{11}e_1+a_{12}e_2+a_{13}e_3,$$

$$v_2=a_{21}e_1+a_{22}e_2+a_{23}e_3,$$

通过直接计算即得

$$v_1 \wedge v_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} e_1 \wedge e_2 + \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} e_2 \wedge e_3 + \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} e_1 \wedge e_3.$$

证毕

最后，在线性空间 $\Lambda^k (k=0, 1, 2, \cdots, n)$ 的基础上构造一个更大的线性空间 Λ，它是 Λ^0，Λ^1，$\cdots$，Λ^n 的直和：

$$\Lambda = \Lambda^0 + \Lambda^1 + \cdots + \Lambda^n,$$

即 Λ 的每一个元素 ω 都可以表示为

$$\omega = \omega_0 + \omega_1 + \cdots + \omega_n,$$

其中 $\omega_i \in \Lambda^i$，并且这一表示是唯一的。在 Λ 内不仅具有线性空间的代数结构，而且还带有外积，称 Λ 是由 V 生成的格拉斯曼(Grassmann)代数，它包含了整个实数域和线性空间 V。作为线性空间，Λ 的基是

$$\begin{array}{ll} 1,\ e_1,\ e_2,\ \cdots,\ e_n, & \\ e_{i_1} \wedge e_{i_2} & (1 \leqslant i_1 < i_2 \leqslant n), \\ e_{i_1} \wedge e_{i_2} \wedge e_{i_3} & (1 \leqslant i_1 < i_2 < i_3 \leqslant n), \\ \cdots\cdots\cdots\cdots, & \\ e_{i_1} \wedge \cdots \wedge e_{i_k} & (1 \leqslant i_1 < \cdots < i_k \leqslant n), \\ \cdots\cdots\cdots\cdots & \\ e_1 \wedge e_2 \wedge \cdots \wedge e_n, & \end{array}$$

它的维数是

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = (1+1)^n = 2^n.$$

行列式

可以用好几种不同的方式来定义行列式，例如一个常用的方式是先给出 $n \times n$ 矩阵，$A = (a_{ij})_{i,j=1,2,\cdots,n}$，然后定义 A 的行列式是一个实数 $\det A$：

$$\det A = \sum_{\binom{12\cdots n}{i_1 i_2 \cdots i_n}} \delta_{i_1 i_2 \cdots i_n} a_{1i_1} a_{2i_2} \cdots a_{ni_n},$$

其中

$$\delta_{i_1\cdots i_n}=\begin{cases}1, & \text{当 } i_1i_2\cdots i_n \text{ 是 } 12\cdots n \text{ 的偶置换；}\\ -1, & \text{当 } i_1i_2\cdots i_n \text{ 是 } 12\cdots n \text{ 的奇置换。}\end{cases}$$

现在，再从外积的角度来定义行列式。设 A 是从线性空间 V 到 V 的一个线性变换，由 A 产生一个从 Λ^n 到 Λ^n 的映射 f，它的定义是

$$f:\ \Lambda^n\to\Lambda^n$$

$$\alpha e_1\wedge\cdots\wedge e_n\mapsto\alpha(Ae_1\wedge Ae_2\wedge\cdots\wedge Ae_n),$$

f 是线性的，这是因为对 Λ^n 中任何两个元素

$$\omega=\alpha e_1\wedge\cdots\wedge e_n,\quad \eta=\beta e_1\wedge\cdots\wedge e_n$$

和任何实数 γ，

$$\begin{aligned}f(\omega+\eta)&=f((\alpha+\beta)e_1\wedge\cdots\wedge e_n)\\&=(\alpha+\beta)(Ae_1\wedge\cdots\wedge Ae_n)\\&=\alpha(Ae_1\wedge\cdots\wedge Ae_n)+\beta(Ae_1\wedge\cdots\wedge Ae_n)\\&=f(\omega)+f(\eta),\\ f(\gamma\omega)&=(\gamma\alpha)(Ae_1\wedge\cdots\wedge Ae_n)\\&=\gamma f(\omega).\end{aligned}$$

而 Λ^n 是 1 维的线性空间，因此从 Λ^n 到 Λ^n 的线性映射只能是“数乘”，即对上述 f，存在一个实数 D，使得对一切 $\omega\in\Lambda^n$，

$$f(\omega)=D\cdot\omega,$$

这个数 D 显然和 $e_1, e_2, \cdots, e_n$ 以及线性变换 A 有关，设 $a_i=Ae_i(i=1, 2, \cdots, n)$，记 $D=\det(a_1, \cdots, a_n)$，即

$$f(\omega)=\det(a_1, \cdots, a_n)\cdot\omega$$

$\det(a_1, \cdots, a_n)$ 称为线性变换 A 的行列式。

这样定义的行列式 $\det(a_1,\cdots,a_n)$ 和原先定义的行列式 $\det A$ 是否相同呢？

性质 3 设 $e_1, e_2, \cdots, e_n$ 是线性空间 V 的一组基，在这组基下线性变换 A 的矩阵表示是 $(a_{ij})_{i,j=1,2,\cdots,n}$，那么

$$\det(a_1, \cdots, a_n)=\sum_{\binom{1\,2\cdots n}{i_1i_2\cdots i_n}}\delta_{i_1\cdots i_n}a_{1i_1}a_{2i_2}\cdots a_{ni_n}=\det A.$$

证明 考察

$$f(e_1\wedge\cdots\wedge e_n)=Ae_1\wedge\cdots\wedge Ae_n,$$

另一方面又有

$$f(e_1\wedge\cdots\wedge e_n)=\det(a_1,\ \cdots,\ a_n)e_1\wedge\cdots\wedge e_n,$$

注意到 $A=(a_{ij})$，再直接计算 $Ae_1\wedge\cdots\wedge Ae_n$，便得出证明。证毕。

习 题

1．设 e_1，e_2，e_3，e_4 是线性空间 X 的基，化简下列 $\omega\wedge\eta$：

(i) $\omega=2e_1+e_2-e_3$，　$\eta=-e_2+5e_3+4e_4$；

(ii) $\omega=-e_1+e_2+2e_3-e_4$，

$$\eta=3e_1\wedge e_2\wedge e_3-e_2\wedge e_3\wedge e_4+6e_1\wedge e_2\wedge e_4;$$

(iii) $\omega=1+2e_3-e_1\wedge e_4+3e_2\wedge e_3\wedge e_4$，

$$\eta=-e_1\wedge e_2\wedge e_3-3e_1\wedge e_3\wedge e_4.$$

2．设 ω 是 k-形式，讨论 $\omega\wedge\omega$。

3．设 V 是线性空间，v_1，v_2，$\cdots$，$v_n\in V$，　证明 v_1，v_2，$\cdots$，v_n 线性无关当且仅当 $v_1\wedge v_2\wedge\cdots\wedge v_n\neq0$。

4．设 V 是 n 维线性空间，ω 是一个 p-形式，$\omega\neq0$，令

$$M_\omega=\{v\in V\mid\omega\wedge v=0\}.$$

证明 M_ω 是 V 的一个子空间，其维数不超过 p，并且其维数等于 p 当且仅当 $\omega=v_1\wedge\cdots\wedge v_p$　(v_1，$\cdots$，v_p 是 V 中的向量)。

5．设 V 是 n 维线性空间，ω 是一个 2-形式，　证明存在 V 的一组基 σ_1，σ_2，$\cdots$，σ_n，使

$$\omega=\sigma_1\wedge\sigma_2+\sigma_3\wedge\sigma_4+\cdots+\sigma_{2k-1}\wedge\sigma_{2k},$$

数 $2k$ 与 ω 有关。

§5.2　微分形式和外微分

本节要讨论的问题是：

(一) 在 n 维欧几里得空间中，如果 f 是可微实值函数，何谓 df？何谓 dx_1，dx_2，$\cdots$，dx_n？应该说，在通常的数学分析中并没有对它给予严格的定义，在那里总是认为 $dx_1=\Delta x_1$，$\cdots$，$dx_n=\Delta x_n$，即自变量的微分等于其改变量，但这是不妥的。本节将指出

$dx_1, \cdots, dx_n$ 是某个线性空间的基，这个线性空间称为微分空间。

（二）在微分空间的基础上，利用 5.1 的方法构造出格拉斯曼代数，其中 k-形式就是 k-微分形式。并利用外积把重积分中一些重要法则化为简单的代数运算。

（三）在微分形式上引进外微分，利用外微分可以得出场论中的三个基本公式，即格林公式、高斯公式和斯托克斯公式，三者有一个统一的形式。其意义并不在于这三个公式的统一，而在于可启发我们把这些公式拓广到更高维的空间中去。

对偶空间

设 X 是一个 n 维实线性空间，φ 是 X 上的一个实线性泛函，这就是说，φ 是从 X 到实数域 R 的一个映射，满足

$$\varphi(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \varphi(x_1) + \alpha_2 \varphi(x_2),$$

其中 $x_1, x_2 \in X$；α_1, α_2 是实数。

记 X^* 是 X 上所有实线性泛函组成的空间，在 X^* 上按通常方法定义"加"和"数乘"两种运算：设 φ_1、$\varphi_2 \in X^*$，α 是实数，定义

$$(\varphi_1 + \varphi_2)(x) = \varphi_1(x) + \varphi_2(x),$$

$$(\alpha\varphi_1)(x) = \alpha\varphi_1(x), \quad \forall x \in X.$$

在 X^* 上装备了上述"加"和"数乘"之后，成为一个(实)线性空间。X^* 称为 X 的代数对偶空间

X^* 的维数是多少？可以选取怎样的元素作为 X^* 的基？下面的定理回答了这两个问题。

定理 1 n 维线性空间 X 的对偶空间 X^* 也是 n 维的线性空间，它的一组基是

$$\varphi_1, \varphi_2, \cdots, \varphi_n,$$

其中线性泛函 φ_i 是这样定义的：设 $e_1, e_2, \cdots, e_n$ 是 X 的一组基，

$$\varphi_i(e_j) = \delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j, \end{cases}$$

$\varphi_1, \varphi_2, \cdots, \varphi_n$ 称为 X^* 内的(关于 $e_1, e_2, \cdots, e_n$ 的)对偶基。

注 对任何$x=x_1e_1+\cdots+x_ne_n\in X$，有

$$\varphi_i(x)=x_1\varphi_i(e_1)+\cdots+x_n\varphi_i(e_n)=x_i,$$

即$\varphi_i(x)$是x的i-坐标。

证明 设f是X上的任意一个线性泛函，$x=x_1e_1+\cdots+x_ne_n$是X内的任意一个元素，那么

$$f(x)=x_1f(e_1)+\cdots+x_nf(e_n),$$

注意到$x_i=\varphi_i(x)\quad(i=1,2,\cdots,n)$，所以

$$f(x)=f(e_1)\varphi_1(x)+\cdots+f(e_n)\varphi_n(x),$$

其中$f(e_1),\cdots,f(e_n)$是仅与f有关(与x无关)的一组数，分别记它们为$\alpha_1,\cdots,\alpha_n$，于是

$$\begin{aligned}f(x)&=\alpha_1\varphi_1(x)+\cdots+\alpha_n\varphi_n(x)\\&=(\alpha_1\varphi_1+\cdots+\alpha_n\varphi_n)(x),\quad\forall x\in X.\end{aligned}$$

这样便证明了

$$f=\alpha_1\varphi_1+\cdots+\alpha_n\varphi_n,$$

即f可以用$\varphi_1,\cdots,\varphi_n$线性表出。

再证明$\varphi_1,\cdots,\varphi_n$线性无关：设存在一组实数$c_1,c_2,\cdots,c_n$，使

$$c_1\varphi_1+\cdots+c_n\varphi_n=0,$$

右端的0是0泛函。将两端作用在e_i上，得

$$(c_1\varphi_1+\cdots+c_n\varphi_n)(e_i)=0,$$

$$c_i=0.$$

所以$\varphi_1,\cdots,\varphi_n$线性无关。证毕。

微分空间

按照第4章中的记号，设T_p是n维微分流形M在点$p\in M$的切空间，点p在坐标邻域U内，U内的局部坐标是$(x_1,x_2,\cdots,x_n)$，T_p的一组基是

$$\left.\frac{\partial}{\partial x_1}\right|_p,\quad\left.\frac{\partial}{\partial x_2}\right|_p,\quad\cdots,\quad\left.\frac{\partial}{\partial x_n}\right|_p,$$

又设ψ是T_p上的一个实线性泛函，即

$$\psi: T_p \to R,$$

并且 ψ 是线性的：

$$\psi(\alpha_1\varphi_1+\alpha_2\varphi_2)=\alpha_1\psi(\varphi_1)+\alpha_2\psi(\varphi_2),$$

其中 φ_1、$\varphi_2 \in T_p$，α_1、α_2 是实数。所有这种 ψ 组成 T_p 的对偶空间 T_p^*，带有通常的“加”和“数乘”，它是一个 n 维的线性空间。

例 1 设 f 是在点 p 可微的函数，由 f 产生 T_p 上的一个泛函 ψ_f 如下：

$$\psi_f: T_p \to R$$

$$\psi_f(\varphi)=\varphi(f),$$

也就是说，设 $\varphi \in T_p$，即

$$\varphi=\alpha_1 \frac{\partial}{\partial x_1}\bigg|_p+\cdots+\alpha_n \frac{\partial}{\partial x_n}\bigg|_p,$$

那么

$$\begin{aligned}\psi_f(\varphi)&=\varphi(f)\\&=\left(\alpha_1 \frac{\partial}{\partial x_1}\bigg|_p+\cdots+\alpha_n \frac{\partial}{\partial x_n}\bigg|_p\right)(f)\\&=\alpha_1 \frac{\partial f}{\partial x_1}(p)+\cdots+\alpha_n \frac{\partial f}{\partial x_n}(p),\end{aligned}$$

它是一个实数。并且容易知道对任何 φ_1、$\varphi_2 \in T_p$ 和任何实数 α，有

$$\begin{aligned}\psi_f(\varphi_1+\varphi_2)&=(\varphi_1+\varphi_2)(f)\\&=\varphi_1(f)+\varphi_2(f)\\&=\psi_f(\varphi_1)+\psi_f(\varphi_2);\\ \psi_f(\alpha\varphi_1)&=(\alpha\varphi_1)(f)\\&=\alpha\varphi_1(f)\\&=\alpha\psi_f(\varphi_1),\end{aligned}$$

所以 ψ_f 是线性的，即 $\psi_f \in T_p^*$。记 $\psi_f=df$，这里的 df 只是一个记号，它是否意味着

$$df=\frac{\partial f}{\partial x_1}(p)dx_1+\cdots+\frac{\partial f}{\partial x_n}(p)dx_n?$$

至少现在还暂时不知道，下面接着继续讨论。

例 2 在例 1 中特别取 $f=x_i$，即对任何 $x=(x_1, \cdots, x_i, \cdots, x_n)$，$f(x)=x_i$，则对任何 $\varphi\in T_p$，

$$\varphi=\alpha_1\frac{\partial}{\partial x_1}\Big|_p+\cdots+\alpha_n\frac{\partial}{\partial x_n}\Big|_p,$$

$$\begin{aligned}\psi_{x_i}(\varphi)&=\varphi(x_i)\\&=\alpha_1\frac{\partial x_i}{\partial x_1}(p)+\cdots+\alpha_n\frac{\partial x_i}{\partial x_n}(p)\\&=\alpha_i\quad(i=1,2,\cdots,n).\end{aligned}$$

如果取 φ 是基向量中的某一个 $\frac{\partial}{\partial x_j}\Big|_p$，那么

$$\psi_{x_i}\left(\frac{\partial}{\partial x_j}\Big|_p\right)=\delta_{ij}=\begin{cases}1, & i=j;\\0, & i\neq j.\end{cases}$$

按照例 1 中的记号，记 $\psi_{x_i}=dx_i$，于是：

(i) $dx_i\left(\frac{\partial}{\partial x_j}\Big|_p\right)=\delta_{ij}$;

(ii) $dx_i(\varphi)=dx_i\left(\alpha_1\frac{\partial}{\partial x_1}\Big|_p+\cdots+\alpha_n\frac{\partial}{\partial x_n}\Big|_p\right)=\alpha_i$.

这时，对例 1 中的 df 就有

$$\begin{aligned}df(\varphi)&=\alpha_1\frac{\partial f}{\partial x_1}(p)+\cdots+\alpha_n\frac{\partial f}{\partial x_n}(p)\\&=dx_1(\varphi)\frac{\partial f}{\partial x_1}(p)+\cdots+dx_n(\varphi)\frac{\partial f}{\partial x_n}(p)\\&=\left(\frac{\partial f}{\partial x_1}(p)dx_1+\cdots+\frac{\partial f}{\partial x_n}(p)dx_n\right)(\varphi),\end{aligned}$$

上式对一切 $\varphi\in T_p$ 均成立，即

$$df=\frac{\partial f}{\partial x_1}(p)dx_1+\cdots+\frac{\partial f}{\partial x_n}(p)dx_n.$$

这和通常数学分析中的全微分是相一致的，但在这里却给出了严格的定义。

由例 1、例 2 和定理 1 立即得到下面的定理 2。

定理 2 切空间 T_p 的对偶空间 T_p^* 是 n 维的线性空间，它的

一组基是

$$dx_1, dx_2, \cdots, dx_n.$$

并且

$$dx_i\left(\frac{\partial}{\partial x_j}\bigg|_p\right) = \delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

即 $dx_1, \cdots, dx_n$ 是一组对偶基。

T_p^* 称为微分流形在点 p 的(局部坐标是$(x_1, \cdots, x_n)$)微分空间。

T_p^* 中元素的一般形式是

$$a_1 dx_1 + \cdots + a_n dx_n,$$

其中 $a_1, \cdots, a_n$ 是定义在点 p 的某邻域内的可微函数。特别当 f 可微时，

$$df = \frac{\partial f}{\partial x_1}(p) dx_1 + \cdots + \frac{\partial f}{\partial x_n}(p) dx_n.$$

微分形式

在微分空间 T_p^*(它是 n 维线性空间)的基础上，利用5.1中引进的方法，由 T_p^* 生成一个格拉斯曼代数 Λ，其中的 k-形式是

$$\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a_{i_1 \cdots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k} \quad (1 \leqslant k \leqslant n),$$

称它是 k-微分形式，或简称为 k-形式，其中 $a_{i_1 \cdots i_k}$ 是可微函数。特别称可微函数 f 是0-形式。

现在给出1-形式的坐标变换。设 ω 是微分流形 M 上的一个1-形式，U 是 M 的一个坐标邻域，局部坐标是$(x_1, \cdots, x_n)$，在 U 内 ω 表示为

$$\omega = a_1(x_1, \cdots, x_n) dx_1 + \cdots + a_n(x_1, \cdots, x_n) dx_n,$$

再设 V 是 M 的另一个坐标邻域，局部坐标是$(y_1, \cdots, y_n)$，在 V 内 ω 表示为

$$\omega = b_1(y_1, \cdots, y_n) dy_1 + \cdots + b_n(y_1, \cdots, y_n) dy_n,$$

在 $U \cap V$ 内(假定它非空)，坐标变换是

$$y_i = y_i(x_1, \cdots, x_n) \quad (i = 1, 2, \cdots, n),$$

那么 a_i 和 b_i 之间有以下变换式：

$$a_i = b_1 \frac{\partial y_1}{\partial x_i} + \cdots + b_n \frac{\partial y_n}{\partial x_i} \quad (i=1,2,\cdots,n).$$

证明 因为

$$dy_i = \frac{\partial y_i}{\partial x_1} dx_1 + \cdots + \frac{\partial y_i}{\partial x_n} dx_n,$$

将它代入 $\omega = b_1 dy_1 + \cdots + b_n dy_n$ 中，得

$$\omega = \sum_{i=1}^{n} b_i \left(\frac{\partial y_i}{\partial x_1} dx_1 + \cdots + \frac{\partial y_i}{\partial x_n} dx_n \right)$$

$$= \sum_{j=1}^{n} \left(b_1 \frac{\partial y_1}{\partial x_j} + \cdots + b_n \frac{\partial y_n}{\partial x_j} \right) dx_j,$$

另一方面

$$\omega = \sum_{j=1}^{n} a_j dx_j,$$

比较 dx_j 的系数即得证明。证毕。

我们再从微分形式和它们之间的外积这一观点来看重积分的变量代换和曲面积分的计算：

对于重积分的变量代换，不失一般性，仅考察二重积分的情形：设二重积分

$$\iint_D f(x,y)\,dxdy,$$

其中 D 是 R^2 中具有零测度边界的有界集，f 在 D 连续，作变量代换

$$x = x(u,v), \quad y = y(u,v), \quad (u,v) \in \sigma.$$

并假定满足变量代换所必须的条件。在这一代换下，$f(x,y)$ 代换为 $f(x(u,v), y(u,v))$；D 将用坐标 (u,v) 表示出来；面积元素 $dxdy$ 将代换为 $\left| \frac{\partial(x,y)}{\partial(u,v)} \right| dudv$，这是大家所熟知的。但有一个问题却常常使不少学生感到迷惑不解——那就是在 $dxdy$ 中，dx 和 dy 分别是 x 和 y 的微分，则在这一代换下有

$$dx = x_u du + x_v dv,$$

$$dy = y_u du + y_v dv,$$

再把 dx 和 dy 相乘，却得不出有意义的结果，那么在面积元素 $dxdy$ 中是不是 dx 乘 dy 呢？从它们的几何直观上看，应该是乘积（微小的矩形的面积），但却不是普通的相乘，而是外积 $dx\wedge dy$，它表示由 dx、dy 所张的平行四边形的有向面积，仅仅因为在笛卡儿直角坐标中它正好是矩形，所以 $dx\wedge dy$ 的绝对值就是 $dxdy$。如果我们把面积元素 $dxdy$ 理解为 $dx\wedge dy$，那么

$$dx\wedge dy=(x_u du+x_v dv)\wedge(y_u du+y_v dv)$$

$$=\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} du\wedge dv.$$

这正是在坐标变换 $x=x(u,v)$，$y=y(u,v)$ 下面积元素的表示式，它是通过形式上的代数运算而获得的，十分简捷。在 n 重 $(n>2)$ 积分的变量代换中也有完全相仿的结果。

再来看曲面积分的计算。设第二类曲面积分是

$$\iint_S f(x,y,z)dxdy,$$

其中 S 是光滑曲面，选定好一侧。设 S 的表达式是

$$x=x(u,v),\quad y=y(u,v),\quad z=z(u,v),\quad (u,v)\in\sigma,$$

f 是 S 上的连续函数。我们假定坐标 u、v 的顺序和坐标 x、y 的顺序是一致的。如果要计算上面的积分，很自然地要用 $f(x(u,v)$、$y(u,v)$、$z(u,v))$ 代替 $f(x,y,z)$，在 S 上的积分当然也将代换为在 $(u,v)\in\sigma$ 上的积分，剩下来的是 $dxdy$，应该把它理解为 $dx\wedge dy$ 还是 $dy\wedge dx$？这和曲面 S 选定哪一侧有关，如果选定 S 所在的侧是上侧（即 S 上的法方向和 z 轴正向的夹角是锐角），那么 $dxdy$ 是 $dx\wedge dy$，这就保证了微小的切平面（它是由 dx 和 dy 张成的）上的法方向 n 和 dx、dy 的排列顺序 dx、dy、n 服从右手法则。如果选定 S 所在的侧是下侧（即 S 上的法方向和 z 轴正向的夹角是钝角），那么 $dxdy$ 是 $dy\wedge dx$。于是

$$dx=x_u du+x_v dv;$$

$$dy=y_u du+y_v dv,$$

$$dx\wedge dy=\frac{\partial(x,y)}{\partial(u,v)}du\wedge dv \qquad (\text{取定} S \text{ 的上侧}),$$

$$dy\wedge dx=-\frac{\partial(x,y)}{\partial(u,v)}du\wedge dv \quad (\text{取定} S \text{ 的下侧}),$$

所以

$$\iint_S f(x,y,z)\,dxdy$$

$$=\pm\iint_\sigma f(x(u,v),y(u,v),z(u,v))\ \frac{\partial(x,y)}{\partial(u,v)}\ dudv,$$

上式右端积分前的符号由选定的 S 所在的侧所确定。这样一来，就把曲面积分的计算通过自然的方式，形式上化为一个简单的代数运算，简捷而明确。

外微分

在所有微分形式组成的格拉斯曼代数 Λ 上引进一个新的映射 d，它把每一个 k-形式($k<n$)映射为$(k+1)$-形式。设

$$\omega=a(x_1,\cdots,x_n)dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

是一个单项的 k-形式，a 是可微（必要时可假定它是任意次可微的）函数，定义

$$d\omega=da\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

$$=\left(\frac{\partial a}{\partial x_1}dx_1+\cdots+\frac{\partial a}{\partial x_n}dx_n\right)\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k},$$

并且还要求 d 是线性的，即 d 是从 Λ 到 Λ 的线性映射，对每一个单项的 k-形式，它的定义如同上面所述的，这一映射 d 称为微分形式上的外微分。

例 3 设局部坐标是(x,y,z)，0-形式 $\omega_0=f(x,y,z)$，则

$$d\omega_0=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dz.$$

又设 1-形式 ω_1 是

$$\omega_1=P(x,y,z)dx+Q(x,y,z)dy+R(x,y,z)dz,$$

则

$$d\omega_1=\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)dy\wedge dz+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)dz\wedge dx$$
$$+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dx\wedge dy.$$

再设 2-形式 ω_2 是

$$\omega_2=A(x,y,z)dy\wedge dz+B(x,y,z)dz\wedge dx$$
$$+C(x,y,z)dx\wedge dy,$$

则

$$d\omega_2=\left(\frac{\partial A}{\partial x}+\frac{\partial B}{\partial y}+\frac{\partial C}{\partial z}\right)dx\wedge dy\wedge dz.$$

此外，对任何 3-形式 ω_3，显然有

$$d\omega_3=0.$$

把上面的结果和场论中的梯度、散度、旋度作一对比是很有意义的。将 dx、dy、dz 分别看作笛卡儿直角坐标系中的 $\boldsymbol{i}$、$\boldsymbol{j}$、$\boldsymbol{k}$；将 $dy\wedge dz$、$dz\wedge dx$、$dx\wedge dy$ 分别看作 $\boldsymbol{j}\times\boldsymbol{k}$、$\boldsymbol{k}\times\boldsymbol{i}$、$\boldsymbol{i}\times\boldsymbol{j}$；将 $dx\wedge dy\wedge dz$ 看作是由 $\boldsymbol{i}$、$\boldsymbol{j}$、$\boldsymbol{k}$ 所张平行 6 面体的有向容积，它等于 1，再令

$$\omega_0=f(x,y,z) \qquad \text{(数量场)},$$
$$\boldsymbol{a}=P(x,y,z)\boldsymbol{i}+Q(x,y,z)\boldsymbol{j}+R(x,y,z)\boldsymbol{k} \qquad \text{(向量场)};$$
$$\boldsymbol{b}=A(x,y,z)\boldsymbol{j}\times\boldsymbol{k}+B(x,y,z)\boldsymbol{k}\times\boldsymbol{i}$$
$$+C(x,y,z)\boldsymbol{i}\times\boldsymbol{j} \qquad \text{(向量场)}.$$

则

$$d\omega_0=\operatorname{grad} f;$$
$$d\boldsymbol{a}=\operatorname{curl}\boldsymbol{a};$$
$$d\boldsymbol{b}=\operatorname{div}\boldsymbol{b}.$$

这告诉我们，场论中的三个基本的量——散度、旋度和梯度——都可以分别看作 0-形式、1-形式和 2-形式的外微分。

再将例 3 和场论中的三个基本公式——格林公式、高斯公式和斯托克斯公式——联系起来，更具有引人入胜的兴趣和产生极有价值的结果。场论中的三个基本公式是（为叙述简单起见，这里

略去公式成立的条件)：

格林公式

$$\iint_D \left(-\frac{\partial P}{\partial y}+\frac{\partial Q}{\partial x}\right)dxdy=\int_{\partial D} Pdx+Qdy;$$

高斯公式

$$\iiint_D \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)dxdydz$$

$$=\iint_{\partial D} Pdydz+Qdzdx+Rdxdy;$$

斯托克斯公式

$$\iint_D \left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)dydz+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)dzdx$$

$$+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dxdy=\int_{\partial D} Pdx+Qdy+Rdz.$$

其中 ∂D 是 D 的边界. 如果令

$$\omega=Pdx+Qdy,$$

则格林公式就是

$$\int_D d\omega=\int_{\partial D}\omega.$$

同样，如果令

$$\omega=Pdy\wedge dz+Qdz\wedge dx+Rdx\wedge dy,$$

或者令

$$\omega=Pdx+Qdy+Rdz,$$

则高斯公式和斯托克斯公式都具有下列形式

$$\int_D d\omega=\int_{\partial D}\omega.$$

这就是场论中三个基本公式的统一形式. 它的作用不仅仅在于这三个公式的统一，而在于启发我们：这一形式可以拓广到更高维的情形中去，在下一章要仔细研究这一问题.

外微分具有下列基本性质：

性质 1 设 ω 是 k-形式，η 是 l-形式，则

$$d(\omega\wedge\eta)=d\omega\wedge\eta+(-1)^k\omega\wedge d\eta.$$

证明 由 d 的线性，不妨假设

$$\omega=a(x_1,\cdots,x_n)dx_{i_1}\wedge\cdots\wedge dx_{i_k};$$
$$\eta=b(x_1,\cdots,x_n)dx_{j_1}\wedge\cdots\wedge dx_{j_l},$$

于是

$$\omega\wedge\eta=abdx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l};$$

$$\begin{aligned}d(\omega\wedge\eta)&=d(ab)\wedge dx_{i_1}\wedge\cdots\wedge dx_{j_l}\\&=\sum_{i=1}^n\left(b\frac{\partial a}{\partial x_i}dx_i+a\frac{\partial b}{\partial x_i}dx_i\right)\wedge dx_{i_1}\wedge\cdots\wedge dx_{j_l}\\&=\sum_{i=1}^n\frac{\partial a}{\partial x_i}dx_i\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge bdx_{j_1}\wedge\cdots\wedge dx_{j_l}\\&\quad+(-1)^k adx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge\sum_{i=1}^n\frac{\partial b}{\partial x_i}dx_i\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l}\\&=d\omega\wedge\eta+(-1)^k\omega\wedge d\eta.\end{aligned}$$ 证毕。

性质 2 对任何微分形式 ω，有

$$d(d\omega)=d^2\omega=0,$$

亦即 $d^2=0$。

证明 由 d 的线性，不妨假设

$$\omega=a(x_1,\cdots,x_n)dx_{i_1}\wedge\cdots\wedge dx_{i_k},$$

则

$$d\omega=\sum_{i=1}^n\frac{\partial a}{\partial x_i}dx_i\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k},$$

$$\begin{aligned}d(d\omega)&=\sum_{i=1}^n d\left(\frac{\partial a}{\partial x_i}\right)\wedge dx_i\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}\\&=\sum_{i=1}^n\sum_{j=1}^n\frac{\partial^2 a}{\partial x_i\partial x_j}dx_j\wedge dx_i\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k},\end{aligned}$$

在 $\sum_{i=1}^n\sum_{j=1}^n\frac{\partial^2 a}{\partial x_i\partial x_j}dx_j\wedge dx_i$ 中包含着两种类型的项：

(i) $\frac{\partial^2 a}{\partial x_i\partial x_i}dx_i\wedge dx_i$;

(ii) $\frac{\partial^2 a}{\partial x_i\partial x_j}(dx_j\wedge dx_i+dx_i\wedge dx_j)$.

无论(i)还是(ii)都等于 0，所以

$$d(d\omega)=0.$$ 证毕。

作为例子，同时也作为对比，利用例 3 后面所说的看法，在 R^3 中将 dx、dy、dz 分别看作坐标轴上的单位向量 i、j、k 等等，设 f 是 0–形式，那么 $d(df)=0$ 就是

$$\operatorname{curl}(\operatorname{grad} f)=0.$$

再设 a 是 1–形式，那么 $d(da)=0$ 就是

$$\operatorname{div}(\operatorname{curl} \boldsymbol{a})=0.$$

这些结论在数学分析的场论中曾叙述过，因此 $d^2\omega=0$ 可以看作是它们的拓广。

现在考虑一个反过来的问题：设定义在 k–形式($k\leqslant n-1$)上的映射 d（暂时不要把它和上面的外微分相混淆），它把每一个 k–形式映射为 $(k+1)$–形式，并且满足

(i) $d(\omega+\eta)=d\omega+d\eta$；

(ii) 设 ω 是 k–形式，η 是 l–形式，那么

$$d(\omega\wedge\eta)=d\omega\wedge\eta+(-1)^k\omega\wedge d\eta;$$

(iii) 对任何 k–形式 ω，

$$d(d\omega)=d^2\omega=0;$$

(iv) 对任何可微函数 f(0–形式)，

$$df=\frac{\partial f}{\partial x_1}dx_1+\cdots+\frac{\partial f}{\partial x_n}dx_n.$$

问：这样的映射 d 是否存在？是否唯一？是否就是上面所给的外微分？下面的定理作了回答。

定理 3 满足上述条件的映射 d 是存在的，唯一的，并且就是微分形式上的外微分。

证明 由上面引进的外微分以及外微分所满足的性质可知道，定理中所说的映射是存在的，例如外微分 d 就是这样的一个映射。

其次证明唯一性如下：

先证明满足上述(i)～(iv)的 d 有

$$d(dx_{i_1}\wedge\cdots\wedge dx_{i_k})=0.$$

利用归纳法，当 $k=1$ 时，对坐标函数 x_{i_1}，由(iv)得

$$d(x_{i_1}) = dx_{i_1},$$

所以 dx_{i_1} 是一个 1-形式，再由(iii)得

$$d(dx_{i_1}) = 0.$$

如果上式对 $k \leqslant p-1$ 成立，当 $k=p$ 时，由(ii)，得

$$\begin{aligned} & d(dx_{i_1} \wedge (dx_{i_2} \wedge \cdots \wedge dx_{i_p})) \\ &= d(dx_{i_1}) \wedge (dx_{i_2} \wedge \cdots \wedge dx_{i_p}) \\ &\quad + (-1) dx_{i_1} \wedge d(dx_{i_2} \wedge \cdots \wedge dx_{i_p}). \end{aligned}$$

由归纳法的假设有

$$d(dx_{i_1}) = 0,$$

$$d(dx_{i_2} \wedge \cdots \wedge dx_{i_p}) = 0,$$

所以

$$d(dx_{i_1} \wedge \cdots \wedge dx_{i_p}) = 0.$$

最后证明由条件 (i)～(iv) 唯一确定了 d。设 ω 是一个 k-形式：

$$\omega = a(x_1, \cdots, x_n) dx_{i_1} \wedge \cdots \wedge dx_{i_k},$$

可以把它看成 0-形式 a 和 k-形式 $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ 的外积，那么

$$\begin{aligned} d\omega &= d(a dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\ &= da \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} + (-1)^0 a d(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\ &= da \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &= \left(\frac{\partial a}{\partial x_1} dx_1 + \cdots + \frac{\partial a}{\partial x_n} dx_n \right) \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \end{aligned}$$

这样便证明了满足(i)～(iv)的 d 就是上面所引进的外微分，并且是唯一确定的。证毕。

闭形式和恰当形式

对微分形式 ω，如果

$$d\omega = 0,$$

称 ω 是闭形式。例如在 R^2 中设

$$\omega = (x^2 + 2xy - y^2) dx + (x^2 - 2xy - y^2) dy,$$

容易验证 $d\omega = 0$，所以它是一个闭形式。此外，由 $d(d\omega) = 0$ 知道，对任何形式 ω，$d\omega$ 是一个闭形式。

对微分形式 ω，如果存在微分形式 η，使

$$\omega = d\eta,$$

称 ω 是恰当形式。例如在 R^3 中设

$$\omega = dx\wedge dy\wedge dz,$$

ω 是一个恰当形式，这是因为若设

$$\eta = \frac{1}{3}\ xdy\wedge dz + \frac{1}{3}\ ydz\wedge dx + \frac{1}{3}\ zdx\wedge dy,$$

那么 $\omega = d\eta$。此外，$d(d\omega) = 0$ 表明每一个恰当形式都是闭形式。

反过来问：闭形式是否一定是恰当形式？

在 R^2 上，我们已经知道下列事实：设 D 是 R^2 内单连通区域（即 D 内任何一条闭曲线可以不经过 D 外的点，而在 D 内连续的收缩为一点），又设函数 P、Q 在 D 内有关于 x 和 y 的连续偏导数，那么下面四个条件是等价的：

(i) 对 D 内任何一条光滑或逐段光滑闭曲线 l，曲线积分

$$\int_l Pdx + Qdy = 0.$$

(ii) D 内的曲线积分

$$\int_l Pdx + Qdy$$

仅与曲线 l 的起点和终点有关，而与路径本身无关。

(iii) 微分式 $Pdx + Qdy$ 是 D 内某个函数 U 的全微分，即

$$dU = Pdx + Qdy,$$

(iv) 在 D 内有

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

成立。

我们再来考察(iii)和(iv)，设

$$\omega = Pdx + Qdy,$$

它是 1-形式，(iii)表明存在 0-形式 $U(x, y)$，使

$$\omega = dU,$$

即 ω 是一个恰当形式，而(iv)又表明

$$d\omega = 0,$$

即 ω 是一个闭形式。因此，在 R^2 内的单连通区域内，1-形式是恰当形式等价于它是闭形式。

这一结论还可以推广到高维空间 R^n 上，设 D 是 R^n 中的一个单连通区域，则在 D 内 1-形式是恰当形式等价于它是闭形式。

但对于 k-形式($k>1$)，情况就变得复杂多了，它对区域 D 提出了更高的要求。设 D 是 R^n 中的一个区域，如果存在点 $x_0\in D$，使对于任何 $x\in D$，连线 xx_0 含在 D 内(图 5-3)，称 D 是关于 x_0 的星形区域。显然，任何凸区域一定是星形区域。

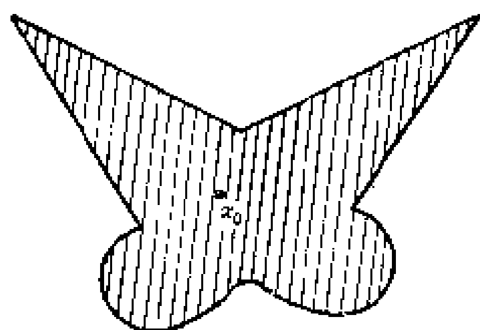


图 5-3

庞卡莱(Poincaré)定理 设 D 是 R^n 中的一个星形区域，则 D 上的任何闭形式一定是恰当形式。

证明 不妨假定 D 是关于原点 O 的星形区域，我们只要证明对任何形式 ω，存在一个线性映射 I，它将 $(k+1)$-形式映射为 k-形式，并且有

$$\omega = I(d\omega) + d(I\omega),$$

如果证明了这一点，那么当 ω 是一个闭形式时，$d\omega = 0$，就有

$$\omega = d(I\omega),$$

这样便证明了结论。

下面证明确实存在这样的映射 I。令

$$\omega = \sum_{1\leqslant i_1<\cdots<i_k\leqslant n} a_{i_1\cdots i_k}(x_1,\cdots,x_n)dx_{i_1}\wedge\cdots\wedge dx_{i_k},$$

它是一个 k-形式，其中 $a_{i_1\cdots i_k}$ 是 D 内连续可微的函数，定义

$$I\omega(x) = \sum_{i_1<\cdots<i_k}\sum_{j=1}^{k}(-1)^{j-1}x_{i_j}\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k},$$

$\widehat{dx_{i_j}}$ 表示在 $dx_{i_1}\wedge\cdots\wedge dx_{i_j}\wedge\cdots\wedge dx_{i_k}$ 中去掉 dx_{i_j}，例如 $dx_1\wedge\widehat{dx_2}\wedge dx_3$ 就是 $dx_1\wedge dx_3$。由于 D 是(关于原点 O 的)星形区域，所以上式右端的积分是有定义的，因此映射 I 有意义。下面验证

$$\omega=I(d\omega)+d(I\omega).$$

对每一项

$$A_j=x_{i_j}\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge\widehat{dx_{i_j}}\wedge\cdots\wedge dx_{i_k}$$

求它的外微分

$$\begin{aligned}dA_j=&\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)\,dt\,dx_{i_j}\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx_{i_j}}\wedge\cdots\wedge dx_{i_k}\\&+x_{i_j}d\left(\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)\,dt\right)\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx_{i_j}}\wedge\cdots\wedge dx_{i_k}\\=&(-1)^{j-1}\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_j}\wedge\cdots\wedge dx_{i_k}\\&+x_{i_j}\sum_{s=1}^n\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt\,dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx_{i_j}}\wedge\cdots\wedge dx_{i_k},\end{aligned}$$

而

$$I\omega=\sum_{i_1<\cdots<i_k}\sum_{j=1}^k(-1)^{j-1}A_j,$$

所以

$$\begin{aligned}d(I\omega)=&k\sum_{i_1<\cdots<i_k}\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}\\&+\sum_{i_1<\cdots<i_k}\sum_{j=1}^k x_{i_j}\sum_{s=1}^n(-1)^{j-1}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt\\&\times dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx_{i_j}}\wedge\cdots\wedge dx_{i_k}.\end{aligned}$$

再来计算 $I(d\omega)$，

$$d\omega=\sum_{i_1<\cdots<i_k}\sum_{s=1}^n\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}\,dx_s\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k},$$

考察 $dx_s\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}$，下标 $i_1,i_2,\cdots,i_k$ 是按大小顺序排列的，但 s 和 $i_1,\cdots,i_k$ 并没有按大小顺序排列。现在将它们按大小顺序重排，那么 s 可能排在奇数位置上，也可能排在偶数位置上。如果 s 在奇数位置上，那么

$$dx_s\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}=dx_{i_1}\wedge\cdots\wedge dx_s\wedge\cdots\wedge dx_{i_k};$$

如果 s 在偶数位置上，那么

$$dx_s\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}=-dx_{i_1}\wedge\cdots\wedge dx_s\wedge\cdots\wedge dx_{i_k}.$$

上面两个式子右端的下标是按大小顺序排列的。

在 $I(d\omega)$ 中，每一项都去掉某个 $dx_{i_j}(1\leqslant j\leqslant k)$ 或者去掉 dx_s。先考虑去掉 dx_s 的项是什么样子的。当 s 在奇数位置上，则这一项是

$$x_s\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k};$$

当 s 在偶数位置上，则这一项是

$$-(-1)x_s\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}.$$

可见不论 s 在奇数位置上还是偶数位置上，去掉 dx_s 的项是相同的。

再考虑去掉 dx_{i_α} 的项。

(i) 当 s 在奇数位置上，如果 $i_\alpha>s$，即 dx_{i_α} 在 dx_s 的后面，那么在 dx_{i_α} 之前（连同 dx_{i_α} 本身）有 α 个 dx_{i_j} 和一个 dx_s，共 $\alpha+1$ 个，所以在 $I(d\omega)$ 中去掉 dx_{i_α} 的项是

$$\begin{aligned}&(-1)^\alpha x_{i_\alpha}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)dt\,dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_\alpha}\wedge\cdots\wedge dx_{i_k}\\&=(-1)^\alpha x_{i_\alpha}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)dt\,dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_\alpha}\wedge\cdots\wedge dx_{i_k}.\end{aligned}$$

如果 $i_\alpha<s$，即 dx_{i_α} 在 dx_s 的前面，这时若将 dx_{i_α} 去掉，那么 s 将处在偶数位置上，因此，去掉 dx_{i_α} 的项是

$$\begin{aligned}&(-1)^{\alpha-1}x_{i_\alpha}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)dt\,dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_\alpha}\wedge\cdots\wedge dx_{i_k}\\&=-(-1)^{\alpha-1}x_{i_\alpha}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)dt\,dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_\alpha}\wedge\cdots\wedge dx_{i_k}.\end{aligned}$$

(ii) 当 s 在偶数位置上，如同(i)的情形那样，可以得到不论 $i_\alpha<s$ 还是 $i_\alpha>s$，去掉 dx_{i_α} 的项是

$$(-1)^\alpha x_{i_\alpha}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)dt\,dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_\alpha}\wedge\cdots\wedge dx_{i_k}.$$

这样一来，就得到

$$I(d\omega)=\sum_{i_1<\cdots<i_k}\sum_{s=1}^{n}x_s\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

$$+\sum_{i_1<\cdots<i_k}\sum_{s=1}^{n}\sum_{\alpha=1}^{k}(-1)^{\alpha}x_{i_\alpha}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt$$

$$\times dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_\alpha}\wedge\cdots\wedge dx_{i_k}.$$

所以

$$d(I\omega)+I(d\omega)=k\sum_{i_1<\cdots<i_k}\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

$$+\sum_{i_1<\cdots<i_k}\sum_{j=1}^{k}\sum_{s=1}^{n}(-1)^{j-1}x_{i_j}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt$$

$$\times dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_j}\wedge\cdots\wedge dx_{i_k}$$

$$+\sum_{i_1<\cdots<i_k}\sum_{s=1}^{n}x_s\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

$$+\sum_{i_1<\cdots<i_k}\sum_{s=1}^{n}\sum_{\alpha=1}^{k}(-1)^{\alpha}x_{i_\alpha}\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt$$

$$\times dx_s\wedge dx_{i_1}\wedge\cdots\wedge\widehat{dx}_{i_\alpha}\wedge\cdots\wedge dx_{i_k}.$$

右端的第二项和第四项正好抵消，于是

$$d(I\omega)+I(d\omega)=k\sum_{i_1<\cdots<i_k}\int_0^1 t^{k-1}a_{i_1\cdots i_k}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

$$+\sum_{i_1<\cdots<i_k}\sum_{s=1}^{n}x_s\int_0^1 t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}(tx)\,dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}.$$

再注意到

$$\frac{d}{dt}\left(t^k a_{i_1\cdots i_k}(tx)\right)=kt^{k-1}a_{i_1\cdots i_k}(tx)$$

$$+\sum_{s=1}^{n}t^k\frac{\partial a_{i_1\cdots i_k}}{\partial x_s}x_s,$$

所以

$$d(I\omega)+I(d\omega)=\sum_{i_1<\cdots<i_k}\int_0^1\frac{d}{dt}\left(t^k a_{i_1\cdots i_k}(tx)\right)dt\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

$$=\sum_{i_1<\cdots<i_k}a_{i_1\cdots i_k}(x)\,dx_{i_1}\wedge\cdots\wedge dx_{i_k}=\omega.\qquad \text{证毕}$$

推论 在 R^n 内任何一个凸区域上，闭形式必是恰当形式。

习 题

1. 在3维微分流形M上，$p\in M$，在点p的坐标邻域内局部坐标是(x, y, z)，p的坐标是$\left(3, \frac{\pi}{3}, -1\right)$，又设$f(x, y, z)=ze^{x}\cos y$，$\varphi\in T_p$，

$$\varphi=3\frac{\partial}{\partial x}\bigg|_p-2\frac{\partial}{\partial y}\bigg|_p+\frac{\partial}{\partial z}\bigg|_p.$$

求$(df)(\varphi)$

2. 在R^3中，设$\varphi_1, \varphi_2, \varphi_3\in T_p$，

$$\varphi_1=3\frac{\partial}{\partial x}-\frac{\partial}{\partial y}, \qquad \varphi_2=\frac{\partial}{\partial y}+\frac{\partial}{\partial z},$$

$$\varphi_3=\frac{\partial}{\partial x}+\frac{\partial}{\partial z}.$$

求$\omega_1, \omega_2, \omega_3\in T_p^*$，使

$$\omega_i(\varphi_j)=\delta_{ij}=\begin{cases}1, & i=j;\\ 0, & i\neq j.\end{cases}$$

并证明T_p中任何元素φ可以表示为

$$\varphi=\omega_1(\varphi)\varphi_1+\omega_2(\varphi)\varphi_2+\omega_3(\varphi)\varphi_3.$$

3. 在R^n中，设$\varphi_1, \cdots, \varphi_n\in T_p$并且线性无关，证明存在$\omega_1, \cdots, \omega_n\in T_p^*$，使

$$\omega_i(\varphi_j)=\delta_{ij},$$

并且对每一个$\varphi\in T_p$，有

$$\varphi=\omega_1(\varphi)\varphi_1+\cdots+\omega_n(\varphi)\varphi_n.$$

又：$\omega_1, \cdots, \omega_n$是线性无关的，因此是$T_p^*$的一组基。

4. 对任何可微函数f和g，证明

$$d(fg)=fdg+gdf.$$

5. 在R^3中求下列形式ω的外微分：

(i) $\omega=(x^2+xy-z)dx+dy-xye^{yz}dx\wedge dz$;

(ii) $\omega=xyzdx+dy\wedge dz+zdx\wedge dy-(x^2+y^2)dx\wedge dy\wedge dz$.

6. 求下列式子的外微分：

(i) $d\omega\wedge\eta-\omega\wedge d\eta$;

(ii) $d\omega\wedge\eta\wedge\zeta+\omega\wedge d\eta\wedge\zeta+\omega\wedge\eta\wedge d\zeta$，其中$\omega$, η都是偶次形式。

7. 在三重积分中，作球面坐标代换$x=r\sin\varphi\cos\theta$, $y=r\sin\varphi\sin\theta$, $z=r\cos\varphi$。求$dx\wedge dy\wedge dz$，并给出积分的代换公式。

8. (i) 当f是0-形式时，已知

$$d(f^2)=2fdf.$$

将这一公式拓广到ω是k-形式(k是偶数)上。

(ii) 当ω是k-形式(k是偶数)，问：$d(\omega\wedge d\omega)=?$

9. 设f是一个复变函数，$dz=dx+idy$，并定义$d(\omega+i\eta)=d\omega+id\eta$。试证明$d(fdz)=0$当且仅当$f$满足柯西-黎曼(Cauchy-Riemann)方程。

10. 证明R^n中的n-形式$\omega=dx_1\wedge\cdots\wedge dx_n$是一个恰当形式。

11. 设ω和η都是闭形式，试证明$\omega\wedge\eta$也是闭形式。

12. 设ω是闭形式，η是恰当形式，试证明$\omega\wedge\eta$是恰当形式。

第6章　流形上的积分·斯托克斯公式

§6.1　流形上的积分

在3维欧几里得空间内有曲线积分、曲面积分和三重积分，它们各自有自身的定义和计算方法。就计算而言，曲线积分是化为定积分来计算的，而曲面积分又是化为二重积分来计算的。在一个高维的欧几里得空间内，例如在R^8内，可以和曲线积分、曲面积分相仿的引进8种积分，倘若都必须各自定义，并给出各自的计算方法，那是非常复杂的事，更不用说在R^8以外还有R^9、R^{10}等等，又该怎么办呢?本节的中心论题是在曲线积分及曲面积分的基础上，利用微分流形的概念，给出一个统一的定义和方法，概括的处理高维空间内的积分学。

关于曲面积分

我们先从读者已经熟悉的曲面积分入手，把微分流形的概念渗透进来，看一看应该怎样合理的定义流形上的积分。

设在R^3内给出了一个第二类曲面积分

$$\iint_S f(x, y, z)\,dxdy,$$

其中S是光滑曲面，f在S上连续，积分沿S上预先给定的一侧(例如上侧或者下侧)。

如果S表示为

$$\psi:\begin{cases} x = x, \\ y = y, \\ z = z(x, y) \quad ((x, y) \in D), \end{cases}$$

那么所给的曲面积分是

$$\iint_S f(x, y, z)dxdy = \iint_D f(x, y, z(x, y))(\pm dx\,dy),$$

右端是一个二重积分，$dxdy$ 前面的符号由 S 上指定的侧所确定，当指定 S 的侧是上侧时，符号取"+"；否则，当指定 S 的侧是下侧时，符号取"−"。

如果 S 表示为

$$\psi: \begin{cases} x = x(u, v), \\ y = y(u, v), \\ z = z(u, v) \quad ((u, v) \in D). \end{cases}$$

则所给的曲面积分就是

$$\iint_S f(x, y, z)dxdy$$
$$= \iint_D f(x(u, v), y(u, v), z(u, v))\left(\pm \frac{\partial(x,y)}{\partial(u,v)}\, dudv\right),$$

同样，右端是一个二重积分，$\frac{\partial(x,y)}{\partial(u,v)}\, dudv$ 前的符号由 S 上指定的侧所确定。

不论 S 怎样表示，上面的两个公式说明了：曲面 S 上的积分可以用一个二重积分来表示．把这一事实说得更仔细一些应该是：设 φ 是 ψ(ψ 是曲面的表示方程)的逆映射，则 φ 就是从 S 到 D 的一个同胚映射，S 是 R^3 中的一个 2 维微分流形，S 上只有一个坐标邻域，它是 S 本身，坐标映射是 φ。而曲面积分

$$\iint_S f(x, y, z)dxdy$$

在这里应该理解为 2-形式 $f(x, y, z)\,dx\wedge dy$ 在 2-维微分流形 S 上的积分，它通过坐标图(S, φ)用欧几里得空间上的一个二重积分来表示，或者说是用后者（二重积分）定义了前者(流形上的积分)。

这一思想可以加以拓广。设 M 是一个 n 维微分流形，并且它

的坐标图册由一个坐标图(M, φ)构成，又设ω是M上的一个n-形式，在坐标$(x_1, x_2, \cdots, x_n)$下，ω表示为

$$\omega = \omega(x_1, \cdots, x_n) dx_1 \wedge \cdots \wedge dx_n,$$

则和上面所说的曲面积分一样，可以把ω在M上的积分$\int_M \omega$定义为

$$\int_M \omega = \int_{\varphi(M)} \omega(x_1, \cdots, x_n) dx_1 \cdots dx_n.$$

右端是欧几里得空间上的n重积分。

但有一个棘手的问题：我们回到2维的曲面上来考虑，如果曲面S不能用一个参式方程表示，也就是说S的坐标图册不是由一个坐标图构成的，而是由(U_i, φ_i)（有限个或无限个）构成的，在每一个坐标邻域U_i内，可以用参数方程φ_i^{-1}表示S的一个局部，按照上面所说，可以给出这一局部上的积分，但如何定义整个S上的积分呢？困难的是在于坐标邻域U_i之间必然会出现许多相交的部分，在这些相交的部分内积分会重复，又如何扣除重复的部分呢？要解决这些问题，关键在于引进1的分解。

1 的分解

设M是一个微分流形，f是定义在M上的一个实值函数，称M的子集

$$\{p \in M \mid f(p) \neq 0\} \text{的闭包}$$

是函数f的支集（或支柱），记为$\text{supp} f$（supp是support的简写）。即

$$\text{supp} f = \overline{\{p \in M \mid f(p) \neq 0\}}.$$

例如在实轴R上，设

$$f(x) = \begin{cases} \sin x, & 0 \leqslant x \leqslant \pi; \\ 0, & x < 0 \text{或} x > \pi. \end{cases}$$

（图6-1），那么

$$\text{supp} f = [0, \pi].$$

在这个例子中，虽然 $\{x\in R|f(x)\neq 0\}=(0,\pi)$，但由于 f 的支集是 f 的所有非 0 点的闭包，所以

$\operatorname{supp} f=[0,\pi]$。



图 6-1

再例如实轴上的函数 $f(x)=x^2-1$，它有两个零点 -1 和 1，但 $\operatorname{supp} f$ 却是整个实轴。

设 M 的坐标邻域是 $U_1, U_2, \cdots, U_i, \cdots$，如果存在 M 上的一列函数 $\pi_1, \pi_2, \cdots, \pi_i\cdots$，使

(i) $\pi_i\in C^\infty$;

(ii) $\pi_i\geqslant 0$，并且 $\operatorname{supp}\pi_i\subset U_i$;

(iii) 对每一个 $p\in M$，只有有限个 i 满足

$$\pi_i(p)\neq 0,$$

(iv) 对每一个 $p\in M$，

$$\sum_i \pi_i(p)=1.$$

(注：由(iii)知道，(iv)中的和实际上是有限项之和)，称 $\{\pi_i\}$ 是 M 上从属于坐标邻域 $\{U_i\}$ 的 1 的 C^∞ 分解。

下面我们考察紧的流形，在一个紧流形 M 中，设它的坐标图册是 $\{(U_\alpha, \varphi_\alpha)\}$，总可以选出有限个坐标邻域 $U_{\alpha_1}, \cdots, U_{\alpha_m}$ 覆盖 M，也就是说，可以用有限个坐标图

$$(U_{\alpha_1}, \varphi_{\alpha_1}), \cdots, (U_{\alpha_m}, \varphi_{\alpha_m})$$

作为 M 的坐标图册。那么是否存在从属于 $U_{\alpha_1}, \cdots, U_{\alpha_m}$ 的 1 的 C^∞ 分解呢？

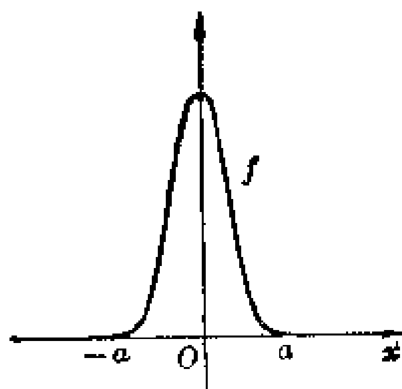


图 6-2

引理 1 设 f 是定义在实轴 R 上的实值函数，

$$f(x)=\begin{cases} e^{-\frac{1}{a^2-x^2}}, & |x|<a \quad (a>0); \\ 0; & |x|\geqslant a. \end{cases}$$

(见图 6-2)，那么

$$f\in C^\infty, \quad 0\leqslant f\leqslant 1,$$

$$\operatorname{supp} f=[-a, a].$$

证明是明显的。

引理 2 设 f 是定义在 n 维欧几里得空间 R^n 上的实值函数，$x=(x_1, x_2, \cdots, x_n)\in R^n$，$f$ 的表示式是

$$f(x)=\begin{cases} e^{-\frac{1}{a^2-|x|^2}}, & |x|=\sqrt{x_1^2+\cdots+x_n^2}<a \ (a>0); \\ 0, & |x|\geqslant a. \end{cases}$$

那么，

$$f\in C^\infty, \ 0\leqslant f\leqslant 1,$$

$$\operatorname{supp} f=\{x\in R^n \mid |x|\leqslant a\}.$$

证明也是明显的。

引理 3 设 M 是 n 维紧微分流形，坐标图册是

$$(U_1, \varphi_1), (U_2, \varphi_2), \cdots, (U_m, \varphi_m).$$

(为方便起见，不妨假定 $\varphi_i(U_i)$ 是 R^n 中以原点为中心、以 r_i 为半径的开球)，则存在开集

$$V_1, V_2, \cdots, V_m,$$

使 (i) $\overline{V}_i\subset U_i$，$\overline{V}_i$ 是 V_i 的闭包；

(ii) $\varphi(V_i)=\{x\in R^n \mid |x|<\delta_i\} \quad (\delta_i<r_i)$；

(iii) $V_1, V_2, \cdots, V_m$ 可覆盖 M，亦即 $(V_1, \varphi_1), (V_2, \varphi_2), \cdots, (V_m, \varphi_m)$ 可作为 M 的坐标图册。

证明 先证明坐标邻域 U_1 可以换为 V_1，即

$$(V_1, \varphi_1), (U_2, \varphi_2), \cdots, (U_m, \varphi_m)$$

可以作为 M 的坐标图册，作

$$A_1=M-(U_2\cup U_3\cup\cdots\cup U_m),$$

因为 $U_2\cup U_3\cup\cdots\cup U_m$ 是开集，所以 A_1 是闭集，并且当 $x\in A_1$ 时，x 一定不属于 $U_2\cup U_3\cup\cdots\cup U_m$，故 $x\in U_1$，即

$$A_1\subset U_1.$$

φ_1 是同胚映射，在 φ_1 的作用下有：

$$\varphi_1(A_1)\subset\varphi_1(U_1),$$

$$\varphi_1(U_1)=B_1=\{x\in R^n \mid |x|<r_1\},$$

$$\varphi_1(A_1) \text{ 是 } B_1 \text{ 内的闭集。}$$

(见图 6-3)，于是存在 δ_1 $(0<\delta_1<r_1)$，使

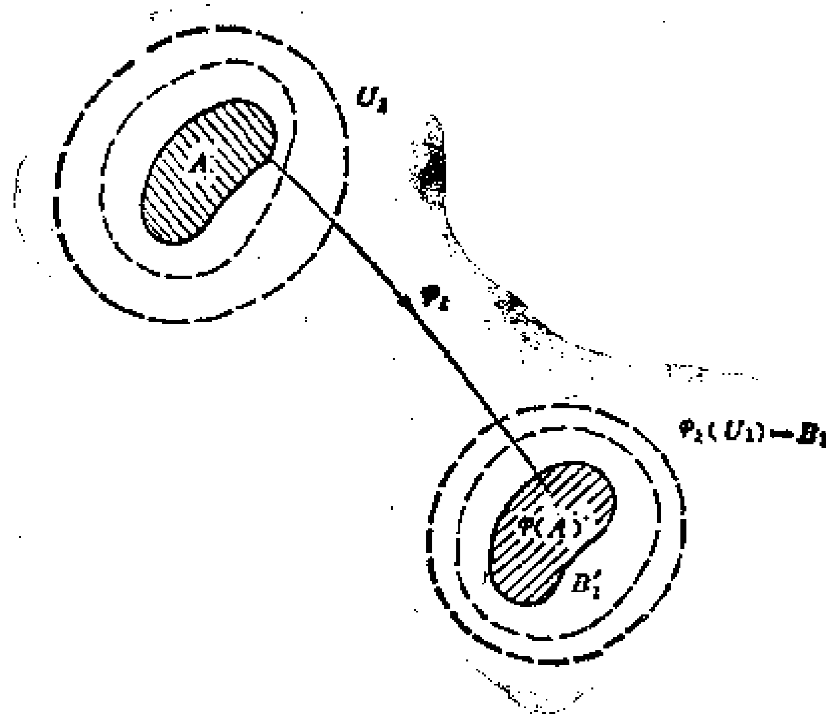


图 6-3

$$\varphi_1(A_1)\subset B_1'=\{x\in R^n \mid |x|<\delta_1\},$$
$$\overline{B}_1'\subset\varphi(U_1)=B_1.$$

记 $$V_1=\varphi_1^{-1}(B_1'),$$

由于 φ_1 是同胚映射，所以

$$A_1\subset V_1,\quad \overline{V}_1\subset U_1,$$

并且 $V_1, U_2, \cdots, U_m$ 覆盖了 M，也就是说

$$(V_1, \varphi_1), (U_2, \varphi_2), \cdots, (U_m, \varphi_m)$$

是M的一个坐标图册。

再用同样的方法可知道，坐标邻域 U_2 可以用 V_2 代替，等等，这样便证明了引理 3。证毕。

定理 1 设M是 n 维紧微分流形，坐标邻域是 $U_1, U_2, \cdots, U_m$，则存在从属于这一坐标邻域的 1 的 C^∞ 分解。

证明 设 $\varphi_1, \varphi_2, \cdots, \varphi_m$ 是相应于 $U_1, U_2, \cdots, U_m$ 的坐标映射，$V_1, V_2, \cdots, V_m$ 是引理 3 所说的可以作为另外一套坐标邻域的开覆盖，$\varphi_i(V_i)$是 R^n 内以原点为中心，以 $\delta_i(\delta_i<r_i)$ 为半径

的开球。

作 R^n 上的实值函数 ψ_i：

$$\psi_i:\ R^n\to R$$

$$\psi_i(x)=\begin{cases}e^{-\frac{1}{\delta_i^2-|x|^2}}, & |x|<\delta_i;\\ 0, & |x|\geqslant\delta_i.\end{cases}$$

由引理2得

$$\psi_i\in C^\infty,\ 0\leqslant\psi_i\leqslant 1,$$

$$\operatorname{supp}\psi_i=\overline{B'_i}\subset B_i=\{x\in R^n\,|\,|x|<r_i\}.$$

再作定义在M上的实值函数 ψ_i^* 如下：

$$\psi_i^*:\ M\to R$$

$$\psi_i^*(p)=\begin{cases}\psi_i\circ\varphi_i(p), & p\in U_i,\\ 0, & p\notin U_i.\end{cases}$$

回忆一下，ψ_i^* 的可微性是指 $\psi_i^*\circ\varphi_i^{-1}=\psi_i\circ\varphi_i\circ\varphi_i^{-1}=\psi_i$ 的可微性，所以有

$$\psi_i^*\in C^\infty,\ \psi_i^*\geqslant 0,$$

$$\operatorname{supp}\psi_i^*=\overline{V}_i\subset U_i,$$

最后对任何 $p\in M$，令

$$\pi_i(p)=\frac{\psi_i^*(p)}{\psi_1^*(p)+\psi_2^*(p)+\cdots+\psi_m^*(p)},$$

因为 $V_1, V_2, \cdots, V_m$ 覆盖了 M，所以右端的分母在任何点 p 都不等于0，此外，容易验证

$$\pi_i\in C^\infty,\ 0\leqslant\pi_i\leqslant 1,$$

$$\operatorname{supp}\pi_i=\operatorname{supp}\psi_i^*=\overline{V}_i\subset U_i,$$

$$\sum_{i=1}^m\pi_i(p)=1,\quad \forall p\in M.$$

证毕

在现代数学的好几个领域内，都有相应的1的分解，只不过随着具体场合的不同，其表达方式亦有所不同罢了。例如在这儿我们要求 $\pi_i\in C^\infty$，但在另外的场合下，有时 π_i 只能是连续的。但主要形式是将单位1分解为满足若干条件的函数之和，而每一个函数只定义在某个局部的范围内(即在它的支集内)，在这个局部

之外函数值是0，因此可以这样说：1的分解在现代数学中的作用是将局部的结果拼接起来。下面就要利用这一思想来定义流形上的积分。

积分

设M是n维定向的紧微分流形，它的坐标图册是

$$(U_1,\ \varphi_1),\ (U_2,\ \varphi_2),\ \cdots,\ (U_m,\ \varphi_m).$$

又设$\pi_1,\ \pi_2,\ \cdots,\ \pi_m$是从属于坐标邻域$\{U_i\}_{i=1,2,\cdots,m}$的1的$C^\infty$分解。再设$\omega$是$M$上的一个$n$-形式。假定在$U_i$内，局部坐标是$(x_1,\ x_2,\ \cdots,\ x_n)$，$\omega$在$U_i$内表示为

$$\omega=a_i(x_1,\ \cdots,\ x_n)dx_1\wedge\cdots\wedge dx_n.$$

定义ω在M上的积分是

$$\int_M\omega=\sum_{i=1}^m\int_M\omega\pi_i.$$

注意到$\mathrm{supp}\pi_i\subset U_i$，所以

$$\int_M\omega=\sum_{i=1}^m\int_{U_i}\omega\pi_i.$$

并且又定义

$$\int_{U_i}\omega\pi_i=\int_{\varphi_i(U_i)}a_i(x_1,\ \cdots,\ x_n)\pi_i(x_1,\ \cdots,\ x_n)dx_1\cdots dx_n,$$

右端是n维欧几里得空间上的一个n重积分。于是

$$\int_M\omega=\sum_{i=1}^m\int_{\varphi_i(U_i)}a_i(x_1,\ \cdots,\ x_n)\pi_i(x_1,\ \cdots,\ x_n)dx_1\cdots dx_n.$$

这一定义是否与1的分解有关呢？作为一个合理的定义当然应该要求它和1的分解无关才行。事实上确实如此。

性质1 上述定义和M上的坐标邻域的选取以及1的C^∞分解无关。

证明 设$U_1,\ U_2,\ \cdots,\ U_m$是M的坐标邻域，$\pi_1,\pi_2,\cdots,\pi_m$是从属于$\{U_i\}_{i=1,2,\cdots,m}$的1的C^∞分解。又设$V_1,\ V_2,\ \cdots,\ V_l$也是M的坐标邻域，$\theta_1,\ \theta_2,\ \cdots,\ \theta_l$是从属于$\{V_j\}_{j=1,2,\cdots,l}$的1的$C^\infty$分解。则不难证明

$$U_i \cap V_j \qquad (i=1, 2, \cdots, m; j=1, 2, \cdots, l).$$

也是M的坐标邻域，并且

$$\pi_i \cdot \theta_j \quad (i=1, 2, \cdots, m; j=1, 2, \cdots, l)$$

是从属于$\{U_i \cap V_j\}$的1的C^∞分解。

设ω是M上的一个n-形式。用坐标邻域$\{U_i\}$覆盖M时，有

$$\int_M \omega = \sum_{i=1}^m \int_{U_i} \omega \pi_i.$$

用坐标邻域$\{V_j\}$覆盖M时，有

$$\int_M \omega = \sum_{j=1}^l \int_{V_j} \omega \theta_j.$$

用坐标邻域$\{U_i \cap V_j\}$覆盖M时，又有

$$\int_M \omega = \sum_{i=1}^m \sum_{j=1}^l \int_{U_i \cap V_j} \omega \pi_i \theta_j.$$

然而

$$\sum_{i=1}^m \int_{U_i} \omega \pi_i = \sum_{i=1}^m \int_{U_i} \omega \pi_i \left(\sum_{j=1}^l \theta_j\right) = \sum_{i=1}^m \sum_{j=1}^l \int_{U_i \cap V_j} \omega \pi_i \theta_j,$$

$$\sum_{j=1}^l \int_{V_j} \omega \theta_j = \sum_{j=1}^l \int_{V_j} \omega \theta_j \left(\sum_{i=1}^m \pi_i\right) = \sum_{j=1}^l \sum_{i=1}^m \int_{U_i \cap V_j} \omega \theta_j \pi_i.$$

这说明不论用怎样的坐标邻域覆盖M以及不论怎样的1的C^∞分解，按上述方式定义的积分其数值是相同的，这就证明了性质1。证毕。

对于非紧的流形可以作相仿的讨论。设M的坐标邻域是U_1，U_2，$\cdots$，U_i，$\cdots$，从属于$\{U_i\}$的1的C^∞分解是π_1，π_2，$\cdots$，π_i，$\cdots$。又设U_i的局部坐标是$(x_1, \cdots, x_n)$，n-形式ω在U_i内表示为

$$\omega = a_i(x_1, \cdots, x_n) dx_1 \wedge \cdots \wedge dx_n,$$

如果

$$\sum_{i=1}^\infty \int_{\varphi_i(U_i)} a_i(x_1, \cdots, x_n) \pi_i(x_1, \cdots, x_n) dx_1 \cdots dx_n$$

存在，就定义它是ω在M上的积分，即

$$\int_M \omega = \sum_{i=1}^\infty \int_{U_i} \omega \pi_i$$

$$= \sum_{i=1}^\infty \int_{\varphi_i(U_i)} a_i(x_1, \cdots, x_n) \pi_i(x_1, \cdots, x_n) dx_1 \cdots dx_n,$$

这里不仔细讨论了.

这样的定义虽然和坐标邻域的选取以及 1 的 C^∞ 分解无关，但在具体计算时总要牵涉到寻找出适当的 1 的 C^∞ 分解，这是相当复杂的事. 不过在许多场合下仍旧可以采取比较方便的方法，例如在 R^3 中求闭球面 S 上的 2-形式 $f(x, y, z)dx\wedge dy$ 的积分，可以把球面 S 分为上半球面和下半球面，分别求出在这两个半球面上的积分，然后再作其和就可以了. 其基本思想是将整个球面沿赤道分开，分成两个开的半球(上半球和下半球)，而去掉一条赤道对积分的值是没有任何影响的.

习　　题

1. 证明从实轴 R 到 R 的函数

$$f(x)=\begin{cases} e^{-\frac{1}{a^2-x^2}}, & |x|<a \quad (a>0); \\ 0; & |x|\geqslant a. \end{cases}$$

是任意阶可微的，即 $f\in C^\infty$. 并求 $\operatorname{supp} f$.

2. 在实轴 R 上，给出从属于 $(n, n+2)$ $(n=0, \pm1, \pm2, \cdots)$ 的 1 的 C^∞ 分解.

3. 在平面 R^2 上，给出从属于 $\{(x, y)\in R^2 | n<x<n+2, m<y<m+2\}$ $(n=0, \pm1, \pm2, \cdots; m=0, \pm1, \pm2, \cdots)$ 的 1 的 C^∞ 分解.

4. 设 M 是 n 维定向的紧微分流形，U 和 V 是它的两个坐标邻域，并且 $U\cap V\neq\phi$，在 U 内局部坐标是 $(x_1, x_2, \cdots, x_n)$；在 V 内局部坐标是 $(y_1, y_2, \cdots, y_n)$，给出积分 $\int_M \omega$ 在 $U\cap V$ 内的坐标变换式. 并利用它来讨论 R^3 中的曲面积分.

§6.2 斯托克斯公式

在上一章引进了外微分之后，我们曾经给出场论中三个基本公式的一个统一形式

$$\int_D d\omega=\int_{\partial D}\omega.$$

当D是平面内的区域，ω是1-形式时，它是格林公式；当D是R^3内的区域，ω是2-形式时，它是高斯公式；当D是R^3中光滑曲面，ω是1-形式时，它是斯托克斯公式。本节的中心是将这一公式拓广到一般的高维空间上，这就是拓广了的（一般的）斯托克斯公式。

在场论的基本公式中，D的边界∂D是有方向的，例如在格林公式中边界∂D是正向(即某人沿边界行走，区域D总是在他的左侧)；在高斯公式中∂D是外侧；在斯托克斯公式中∂D的方向和D的方向之间服从右手法则。为了获得一般的斯托克斯公式，我们首先研究n维欧几里得空间中的单形和正方形的边界如何给定方向，然后再证明一般的斯托克斯公式。

单形和链

单形是最简单的几何体，例如在3维欧几里得空间内，最简单的3维几何体是四面体(图6-4)，它的边界由四个三角形组成，这就是一个3-单形。现在将这一概念拓广如下：设$x_0, x_1, \cdots, x_k$是R^n中的$k+1$个点，这里$k\leqslant n$，并且向量

$$v_1=x_1-x_0,\quad v_2=x_2-x_0,\ \cdots,\ v_k=x_k-x_0$$

是线性无关的。令

$$x=t_0x_0+t_1x_1+\cdots+t_kx_k,$$

其中$t_0+t_1+\cdots+t_k=1$，$t_i\geqslant 0(i=0, 1, 2, \cdots, k)$。所有这种$x$的全体组成$R^n$中的一个子集，称它是$k$-单形，记为$(x_0, x_1, \cdots, x_k)$，并称$a_i$是这个$k$-单形的$i$-顶点，$(i=0, 1, 2, \cdots, k)$。

例如0-单形就是一个单独的点$\{x_0\}$；1-单形是直线上的一个

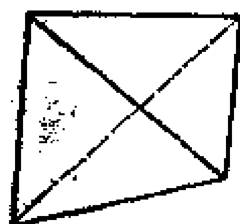


图 6-4



图 6-5

闭的线段(x_0, x_1)；2-单形是三角形(x_0, x_1, x_2)（见图6-5）；3-单形如图6-4所示，等等，它们是代数拓扑中最简单最基本的几何体，也是代数拓扑中的一个最基本的概念。

对一个k-单形$(x_0, x_1, \cdots, x_k)$，它的边界是由$k+1$个$(k-1)$-单形组成的，这些$(k-1)$-单形是

$$(x_0, x_1, \cdots, \hat{x}_i, \cdots, x_k) \quad (i=0, 1, \cdots, k).$$

$\hat{x}_i$表示将x_i去掉，即

$$(x_0, x_1, \cdots, \hat{x}_i \cdots, x_k) = (x_0, x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_k).$$

规定S的边界∂S是

$$\begin{aligned}\partial S &= \partial(x_0, x_1, \cdots, x_k) \\ &= \sum_{i=0}^{k}(-1)^i(x_0, x_1, \cdots, \hat{x}_i, \cdots, x_k),\end{aligned}$$

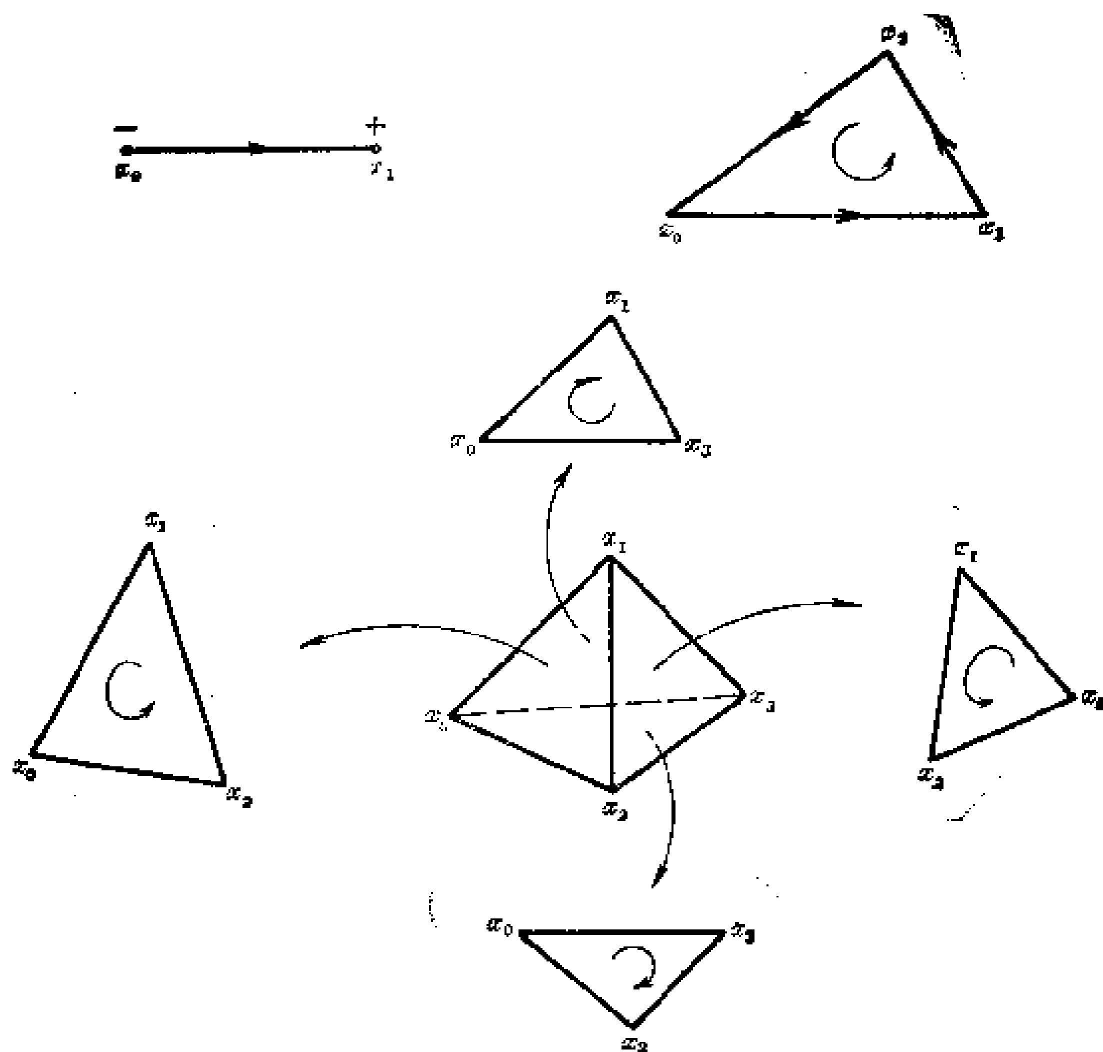


图 6-6

称∂是边界算子，S的边界∂S是由$k+1$个带有符号的$(k-1)$-单形组合起来的。

例如1-单形(x_0, x_1)，2-单形(x_0, x_1, x_2)，3-单形(x_0, x_1, x_2, x_3)的边界分别是

$$\partial(x_0, x_1) = (x_1) - (x_0);$$

$$\partial(x_0, x_1, x_2) = (x_1, x_2) - (x_0, x_2) + (x_0, x_1);$$

$$\partial(x_0, x_1, x_2, x_3) = (x_1, x_2, x_3) - (x_0, x_2, x_3) + (x_0, x_1, x_3) - (x_0, x_1, x_2).$$

图6-6画出了它们的边界，并用箭头表示了边界的方向。

图6-6告诉我们，在平面上三角形区域的边界方向是逆时针方向(即正向)，R^3内四面体的边界的方向是外法向，这些都和数学分析中的规定相吻合。但读者可能会产生一个问题：以2-单形而论，(x_0, x_1, x_2)也可以画成图6-7中的样子，这时它的边界的方向岂不是成了顺时针方向？事实上，在欧几里得空间内还应该有一个约定，当我们画出三角形(x_0, x_1, x_2)时，我们要求向量

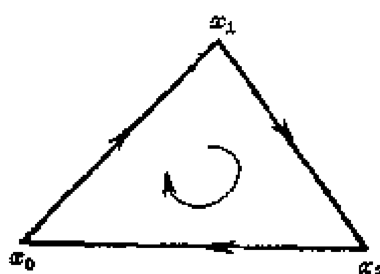


图 6-7

$$v_1 = x_1 - x_0 \text{ 与 } v_2 = x_2 - x_0$$

的顺序安排和平面上笛卡儿坐标x、y轴的顺序安排是一致的，也就是说，从v_1旋转到v_2(旋转角度应在180°之内)的方向和从x轴旋转到y轴的方向是一致的，这时(x_0, x_1, x_2)的图形就是图6-6所画出的。3-单形(x_0, x_1, x_2, x_3)的情形相仿，要求向量

$$v_1 = x_1 - x_0, \quad v_2 = x_2 - x_0, \quad v_3 = x_3 - x_0$$

之间服从右手法则，如同x轴，y轴，z轴那样。

设$S_i (i = 1, 2, \cdots, m)$是$k$-单形，$a_i$是整数，称下列形式的和

$$C = \sum_{i=1}^{m} a_i S_i$$

是k-链。例如3-单形的边界就是一个2-链。

定义k-链C的边界∂C是

$$\partial C = \sum_{i=1}^{m} a_i \partial S_i.$$

边界算子有一个基本性质：对任何链 C，有

$$\partial^2 C = \partial(\partial C) = 0,$$

即 $\partial^2 = 0$。这只要对单形进行验证就可以了。为简单起见，仅验证 2-单形 (x_0, x_1, x_2)。

$$\partial(x_0, x_1, x_2) = (x_1, x_2) - (x_0, x_2) + (x_0, x_1);$$

$$\begin{aligned}\partial(\partial(x_0, x_1, x_2)) &= \partial(x_1, x_2) - \partial(x_0, x_2) + \partial(x_0, x_1) \\ &= [(x_2) - (x_1)] - [(x_2) - (x_0)] \\ &\quad + [(x_1) - (x_0)] = 0.\end{aligned}$$

请读者验证 3-单形，以及一般的 k-单形。

一个 k-形式 ω 在 k-链 C

$$C = \sum_{i=1}^{m} a_i S_i$$

上的积分定义为

$$\int_C \omega = \int_{\sum\limits_{i=1}^{m} a_i S_i} \omega = \sum_{i=1}^{m} a_i \int_{S_i} \omega.$$

n 维单位正方形的边界

现在讨论 R^n 内单位正方形的边界：设

$$I^n = \{(x_1, x_2, \cdots, x_n) \in R^n \mid 0 \leqslant x_1, x_2, \cdots, x_n \leqslant 1\},$$

它是 R^n 内的 n 维单位正方形，它的边界是由 $2n$ 个面组成的，这 $2n$ 个面是

$$I^n_{(i,0)} = \{(x_1, \cdots, x_{i-1}, \overset{i}{0}, x_{i+1}, \cdots, x_n) \mid 0 \leqslant x_1, \cdots, x_n \leqslant 1\};$$

$$I^n_{(i,1)} = \{(x_1, \cdots, x_{i-1}, \overset{i}{1}, x_{i+1}, \cdots, x_n) \mid 0 \leqslant x_1, \cdots, x_n \leqslant 1\}$$

$$(i = 1, 2, \cdots, n).$$

称 $I^n_{(i,0)}$ 和 $I^n_{(i,1)}$ 分别是 I^n 的 $(i, 0)$ 面和 $(i, 1)$ 面。它们本身又都是 R^{n-1} 内的 $n-1$ 维单位正方形。

规定 I^n 的有向边界 ∂I^n 是

$$\partial I^n = \sum_{i=1}^{n} \sum_{\alpha=0,1} (-1)^{i+\alpha} I^n(i, \alpha).$$

它是 $2n$ 个带有符号的面的组合．这一规定是从单形的边界推断出来的，单位正方形可以看作是由单形拼接起来的，为了简单起见，我们只考察 1 维（下面的例 1）和 2 维（下面的例 2）的情形．

例 1 在实轴上，单位正方形就是区间 $[0, 1]$，即

$$I^1 = \{x_1 \in R \mid 0 \leqslant x_1 \leqslant 1\}.$$

它的边界（图 6-8）

$$\partial I^1 = I^1(1, 1) - I^1(1, 0) = \{1\} - \{0\}.$$

这也正是 1-单形 $(0, 1)$ 的边界．

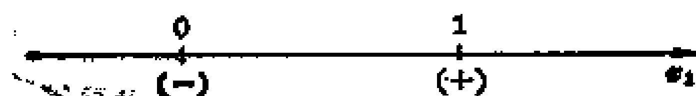


图 6-8

例 2 在平面 R^2 上，

$$I^2 = \{(x_1, x_2) \in R^2 \mid 0 \leqslant x_1, x_2 \leqslant 1\},$$

I^2 的边界 ∂I^2 是（见图 6-9）

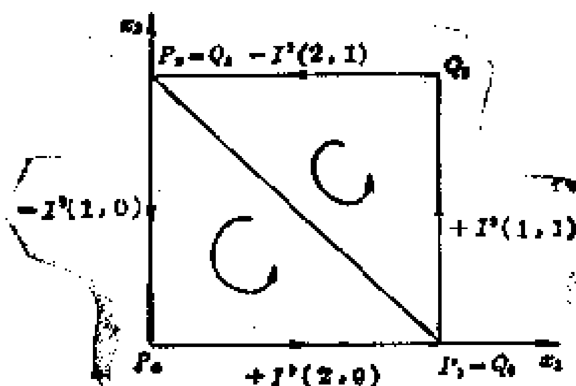


图 6-9

$$\partial I^2 = I^2(2, 0) + I^2(1, 1) - I^2(2, 1) - I^2(1, 0).$$

另一方面可以把 I^2 看作两个 2-单形 (P_0, P_1, P_2) 和 (Q_0, Q_1, Q_2) 的组合 $(P_0, P_1, P_2) + (Q_0, Q_1, Q_2)$（见图 6-9），则

$$\begin{aligned}\partial I^2 &= \partial(P_0, P_1, P_2) + \partial(Q_0, Q_1, Q_2) \\ &= (P_1, P_2) - (P_0, P_2) + (P_0, P_1)\end{aligned}$$

$$+(Q_1, Q_2)-(Q_0, Q_2)+(Q_0, Q_1)$$
$$=(P_0, P_1)+(Q_0, Q_1)+(Q_1, Q_2)-(P_0, P_2).$$
两者是一致的。

图 6-9 表明，2 维单位正方形的边界的方向正是数学分析中所说的正向，也就是逆时针方向(或人沿边界行走时区域总在其左侧)。

下面计算 $\partial(\partial I^2)$：

$$\begin{aligned}\partial^2 I^2 &= \partial(\partial I^2)\\ &= \partial I^2(2, 0)+\partial I^2(1, 1)-\partial I^2(2, 1)-\partial I^2(1, 0)\\ &= (\{P_1\}-\{P_0\})+(\{Q_1\}-\{Q_0\})+(\{Q_2\}-\{Q_1\})\\ &\quad +(\{P_0\}-\{P_2\})=0.\end{aligned}$$

这一事实具有一般性，即对任何 n 维单位正方形 I^n，总有

$$\partial^2 I^n=\partial(\partial I^n)=0.$$

也就是说

$$\partial^2=0.$$

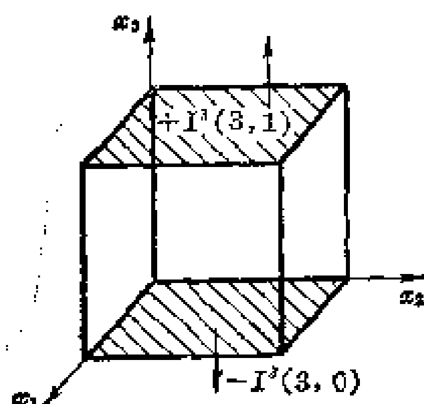
例 3 在 R^3 中

$$I^3=\{(x_1, x_2, x_3)\in R^3\ 0\leqslant x_1, x_2, x_3\leqslant 1\},$$

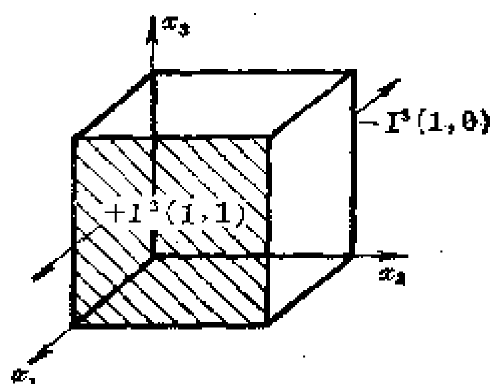
它的边界是

$$\partial I^3=-I^3(1, 0)+I^3(1, 1)+I^3(2, 0)$$
$$-I^3(2, 1)-I^3(3, 0)+I^3(3, 1).$$

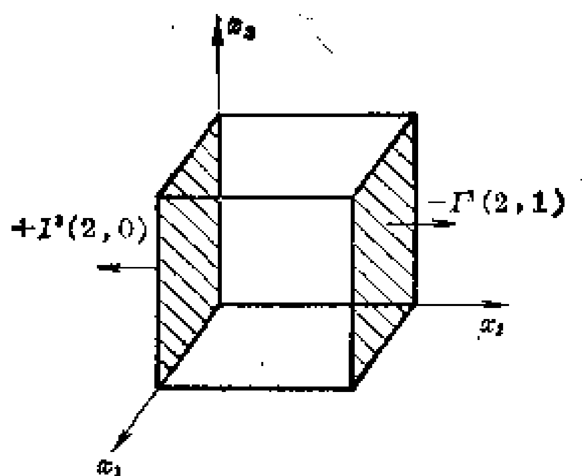
图 6-10 画出了 I^3 的各个边界面和它们的方向。



对 $I^3(3, 0)$和$I^3(3, 1)$而言，坐标是 $x_1x_2\hat{x}_3$. ($\hat{x}_3$ 表示去掉 x_3)



对 $I^3(1, 0)$和 $I^3(1, 1)$而言，坐标是 $\hat{x}_1x_2x_3$. ($\hat{x}_1$ 表示去掉 x_1)



对 $I^3(2, 0)$和$I^3(2, 1)$而言，坐标是 $x_1\hat{x}_2x_3$. ($\hat{x}_2$表示去掉 x_2)

图 6-10

图中表明 R^3 内单位正方形的边界是外侧，这和数学分析中的规定是完全吻合的。

读者不难求出

$$\partial^2 I^3 = \partial(\partial I^3) = 0_\circ$$

积分区域的边界

设M是 n 维微分流形，D是M的一个子集，为了研究在D上的积分，就必须对D的边界作一个约定(或限定)，一般说来，按边界的定义，D的边界是由这种点P组成的，在含有P的任何开集内，既有D内的点又有 $M-D$(D 的补集)内的点，根据这一定义，D的边界可能很复杂，以致于很难处理．在应用上，对D的边界总是作如下的一些约定：

设$U_1, U_2, \cdots, U_i, \cdots$是$M$的坐标邻域，相应的坐标映射是 $\varphi_1, \varphi_2, \cdots, \varphi_i, \cdots$，$D$是$M$的一个子集，$D$的边界记为$\partial D$。如果某一个坐标邻域$U_k$ 和 ∂D 相交，则要求(图 6-11)：

(i) $\varphi_k(U_k) = \{(x_1, \cdots, x_n) \in R^n | 0 < x_1, \cdots, x_{n-1} < 1, -1 < x_n < 1\}$,

即 $\varphi_k(U_k)$是 n 维矩形(注：这一要求非本质的，作为微分流形，为方便起见，可以假设 U_k 和 R^n 内的 n 维矩形同胚)；

(ii) $\varphi_k(\partial D \cap \bar{U}_k)$

$$=\{(x_1, \cdots, x_{n-1}, 0) \in R^n | 0 < x_1, \cdots, x_{n-1} < 1\};$$

(iii) $\varphi_k((\partial D \cup D) \cap \bar{U}_k)$

$$=\{(x_1, \cdots, x_n) \in R^n | 0 < x_1, \cdots, x_{n-1} < 1, 0 \leqslant x_n < 1\}.$$

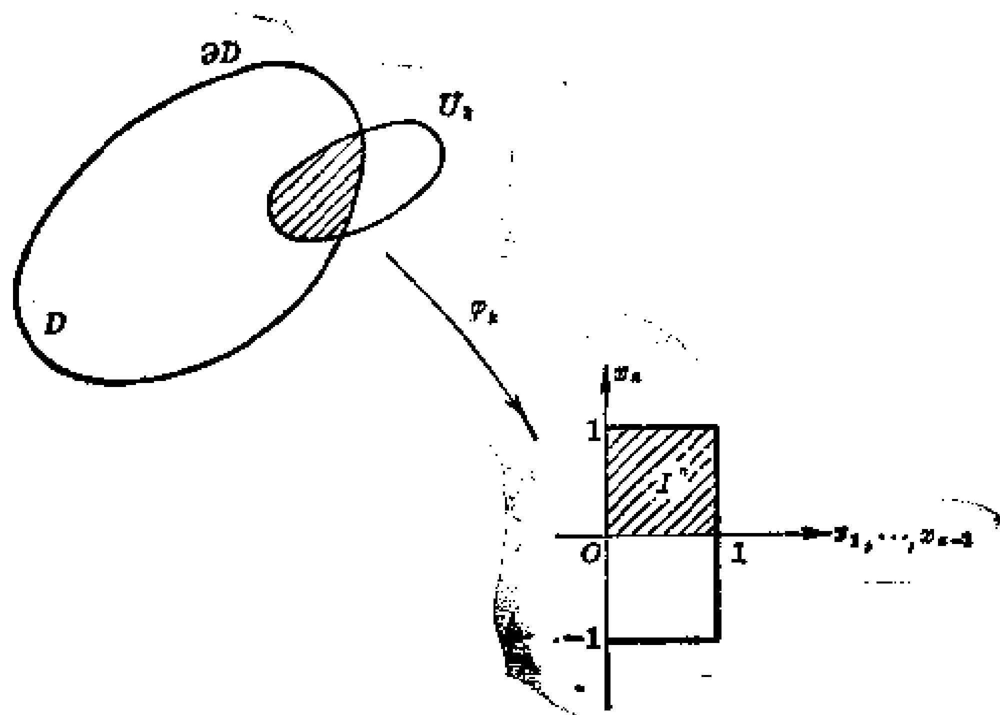


图 6-11

其本质的要求是将位于 U_k 内的那部分边界在坐标映射 φ_k 的作用下映射为 R^n 中的 n-1 维超平面 $x_n=0(0<x_1, \cdots, x_{n-1}<1)$，并把位于 U_k 内的子集 D 的部分映射为上半空间 $x_n>0$ 的一个 n 维单位正方形

$$I^n=\{(x_1, \cdots, x_n) \in R^n | 0 < x_1, \cdots, x_n < 1\}.$$

并且由 I^n 的边界方向来规定 ∂D 的方向。

满足上述条件(i)～(iii)的子集 D 就是我们要研究的对象。

斯托克斯公式

现在给出一般的斯托克斯公式如下：

斯托克斯定理 设 M 是一个 n 维微分流形，并且是紧的和定

向的。D是M内的一个子集，D和它的边界满足上述条件（i）～（iii），则对任何$(n-1)$-形式ω，有

$$\int_D d\omega = \int_{\partial D} \omega$$

（这就是一般的**斯托克斯公式**）。

在证明斯托克斯定理之前先证明一个引理：

引理 如果对任何给定的坐标邻域U_i以及任意一个满足

$$\operatorname{supp} \omega \subset U_i$$

的$(n-1)$-形式ω，有

$$\int_D d\omega = \int_{\partial D} \omega$$

成立。则对每一个$(n-1)$-形式ω，有

$$\int_D d\omega = \int_{\partial D} \omega$$

成立。换句话说，当斯托克斯公式局部成立时，在整体上也成立。

证明 设M是由坐标邻域U_1，U_2，$\cdots$，U_m覆盖的，π_1，π_2，$\cdots\pi_m$是从属于$\{U_i\}$的1的C^∞分解，将π_1，π_2，$\cdots$，π_m限制在∂D上，它就是∂D上从属于$\{\partial D\cap U_i\}$的1的C^∞分解。

又设ω是任意一个$(n-1)$-形式，那么

$$\int_{\partial D} \omega = \sum_{i=1}^{m} \int_{\partial D} \pi_i \omega,$$

因为

$$\operatorname{supp} \pi_i \omega \subset U_i,$$

故由引理的条件知道

$$\int_D d(\pi_i \omega) = \int_{\partial D} \pi_i \omega,$$

于是

$$\int_{\partial D} \omega = \sum_{i=1}^{m} \int_D d(\pi_i \omega).$$

另一方面

$$\int_D d\omega = \sum_{i=1}^{m} \int_D \pi_i d\omega,$$

问题化为只要证明

$$\sum_{i=1}^{m} d(\pi_i \omega) = \sum_{i=1}^{m} \pi_i d\omega.$$

由于 π_i 是一个 0-形式，ω 是一个$(n-1)$-形式，$\pi_i\omega$ 就是 $\pi_i\wedge\omega$，利用外积的外微分公式(5.2 性质 1)，有

$$d(\pi_i\omega)=d\pi_i\wedge\omega+\pi_i d\omega,$$

所以

$$\sum_{i=1}^{m}d(\pi_i\omega)=\left(\sum_{i=1}^{m}d\pi_i\right)\wedge\omega+\sum_{i=1}^{m}\pi_i d\omega$$

$$=d\left(\sum_{i=1}^{m}\pi_i\right)\wedge\omega+\sum_{i=1}^{m}\pi_i d\omega,$$

但 $\pi_1, \pi_2, \cdots, \pi_m$是 1 的 C^∞ 分解：$\sum_{i=1}^{m}\pi_i=1$，所以

$$d\left(\sum_{i=1}^{m}\pi_i\right)=0,$$

这样便证明了结论。证毕。

再证明斯托克斯定理如下：

证明 只要证明引理的条件满足就可以了，设 ω 是任意一个$(n-1)$-形式，并且

$$\operatorname{supp}\omega\subset U,$$

U 是M的某一个坐标邻域，我们要证明的是

$$\int_D d\omega=\int_{\partial D}\omega.$$

设相应于坐标邻域U的坐标映射是 φ，局部坐标是 $(x_1, x_2, \cdots, x_n)$。这时有两种情形须要考虑：

情形 1　坐标邻域U和D的边界 ∂D 不相交，U在D的内部(图 6-12)。在坐标映射 φ 的作用下，坐标邻域U和 n 维单位正方形 $\varphi(U)=\{(x_1, \cdots, x_n)\in R^n|0<x_1, \cdots, x_n<1\}$同胚。

情形 2　坐标邻域U和D的边界 ∂D 相交(图 6-11)，如同前面所说的那样：

$$\varphi(U)=\{(x_1, \cdots, x_n)\in R^n|0<x_1, \cdots, x_{n-1}<1,$$
$$-1<x_n<1\},$$

$$\varphi((D\cup\partial D)\cap U)$$
$$=\{(x_1, \cdots, x_n)\in R^n|0<x_1, \cdots, x_{n-1}<1, 0\leqslant x_n<1\},$$

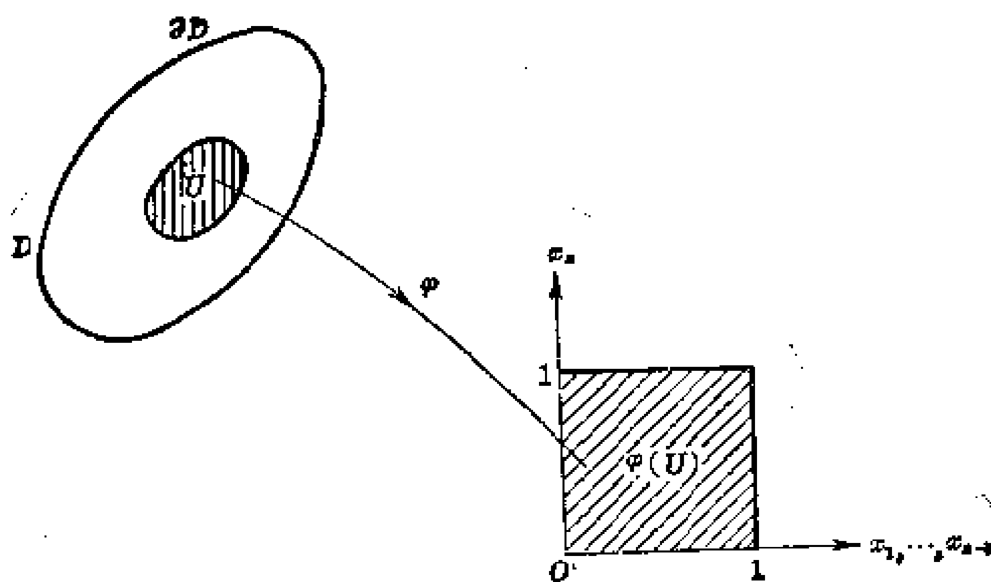


图 6-12

$$
\begin{aligned}
&\varphi(\partial D\cap U)\\
&=\{(x_1,\ \cdots,\ x_{n-1},\ 0)\in R^n\,|\,0<x_1,\ \cdots,\ x_{n-1}<1\}\\
&=I^n(n,\ 0).
\end{aligned}
$$

后者是 R^n 中 n 维单位正方形 I^n 的$(n,\ 0)$面，按边界的符号规则，它的符号是$(-1)^n$。

设$(n-1)$-形式 ω 是(在坐标邻域U内用局部坐标$(x_1,\ x_2,\ \cdots,\ x_n)$表示出来)

$$\omega=\sum_{j=1}^{n}a_j(x_1,\ \cdots,\ x_n)dx_1\wedge\cdots\wedge\widehat{dx_j}\wedge\cdots\wedge dx_n,$$

其中 $\widehat{dx_j}$ 表示去掉 dx_j。并由假设有

$$\operatorname{supp}a_j\subset\varphi(U).$$

这时

$$
\begin{aligned}
d\omega&=\sum_{j=1}^{n}(da_j)\wedge dx_1\wedge\cdots\wedge\widehat{dx_j}\wedge\cdots\wedge dx_n\\
&=\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial a_j}{\partial x_i}\,dx_i\wedge dx_1\wedge\cdots\wedge\widehat{dx_j}\wedge\cdots\wedge dx_n,
\end{aligned}
$$

在这个和式中，对 i 来说，当 $i\neq j$ 时，

$$dx_i\wedge dx_1\wedge\cdots\wedge\widehat{dx_j}\wedge\cdots\wedge dx_n=0,$$

所以只有当 $i=j$ 的项可能不是 0，于是

$$d\omega=\sum_{j=1}^{n}\frac{\partial a_j}{\partial x_j}dx_j\wedge dx_1\wedge\cdots\wedge\widehat{dx_j}\wedge\cdots\wedge dx_n$$

$$=\sum_{j=1}^{n}(-1)^{j-1}\frac{\partial a_j}{\partial x_j}dx_1\wedge\cdots\wedge dx_j\wedge\cdots\wedge dx_n,$$

这时 $$\int_D d\omega=\int_{U\cap D}d\omega$$

$$=\sum_{j=1}^{n}(-1)^{j-1}\underbrace{\int_0^1\cdots\int_0^1}_{n}\frac{\partial a_j}{\partial x_j}dx_1\cdots dx_n,$$

将它先对 x_j 积分：

$$\int_D d\omega=\sum_{j=1}^{n}(-1)^{j-1}\underbrace{\int_0^1\cdots\int_0^1}_{n-1}[a_j(x_1,\cdots,x_{j-1},1,x_{j+1},\cdots,x_n)$$

$$-a_j(x_1,\cdots,x_{j-1},0,x_{j+1},\cdots,x_n)]dx_1\cdots\widehat{dx_j}\cdots dx_n,$$

下面分情形 1 和情形 2 来分别考虑：

在情形 1 时，因为

$\operatorname{supp}a_j\subset\varphi(\bar{U})=I^n=\{(x_1,\cdots,x_n)\in R^n|0<x_1,\cdots,x_n<1\}$。

而点 $(x_1,\cdots,x_{j-1},1,x_{j+1},\cdots,x_n)$ 和点 $(x_1,\cdots,x_{j-1},0,x_{j+1},\cdots,x_n)$（其中 $0<x_1,\cdots,x_{j-1},x_{j+1},\cdots,x_n<1$；$j=1,2,\cdots,n$）都在 I^n 的边界上，所以

$$a_j(x_1,\cdots,x_{j-1},1,x_{j+1},\cdots,x_n)=0,$$

$$a_j(x_1,\cdots,x_{j-1},0,x_{j+1},\cdots,x_n)=0\quad(j=1,2,\cdots,n).$$

于是 $$\int_D d\omega=0.$$

另外又因为

$$\operatorname{supp}\omega\subset U.$$

而 U 在 D 的内部，显然有

$$\int_{\partial D}\omega=0.$$

这样便证明了在情形 1 时

$$\int_D d\omega=\int_{\partial D}\omega.$$

再考察情形2，同样，点$(x_1, \cdots, x_{j-1}, 1, x_{j+1}, \cdots, x_n)$ $(j=1, 2, \cdots, n)$ 和点 $(x_1, \cdots, x_{j-1}, 0, x_{j+1}, \cdots, x_n)$ (但这里的 $j=1, 2, \cdots, n-1$)都在

$\varphi(U)=\{(x_1, \cdots, x_n)\in R^n \mid 0<x_1, \cdots, x_{n-1}<1, -1<x_n<1\}$

的边界上，但点$(x_1, \cdots, x_{n-1}, 0)$是在$\varphi(U)$的内部，因此

$a_j(x_1, \cdots, x_{j-1}, 1, x_{j+1}, \cdots, x_n)=0 \quad (j=1, 2, \cdots, n);$

$a_j(x_1, \cdots, x_{j-1}, 0, x_{j+1}, \cdots, x_n)=0 \quad (j=1, 2, \cdots, n-1);$

$a_j(x_1, \cdots, x_{n-1}, 0)$可能不是0。

代入刚才给出的$\int_D d\omega$中，得

$$\int_D d\omega=(-1)^n\underbrace{\int_0^1\cdots\int_0^1}_{n-1}a_n(x_1, \cdots, x_{n-1}, 0)dx_1\cdots dx_{n-1}.$$

另一方面，如果将ω限制在$\partial D\cap U$上，在局部坐标的表示下，即限制在$x_n=0$, $0<x_1, \cdots, x_{n-1}<1$上，所以$dx_n=0$，这时

$$\begin{aligned}\omega&=\sum_{j=1}^n a_j(x_1, \cdots, x_{n-1}, 0)dx_1\wedge\cdots\wedge\widehat{dx_j}\wedge\cdots\wedge dx_n\\&=a_n(x_1, \cdots, x_{n-1}, 0)dx_1\wedge\cdots\wedge dx_{n-1},\end{aligned}$$

所以
$$\begin{aligned}\int_{\partial D}\omega&=\int_{\partial D\cap U}\omega\\&=(-1)^n\underbrace{\int_0^1\cdots\int_0^1}_{n-1}a_n(x_1, \cdots, x_{n-1}, 0)dx_1\cdots dx_{n-1}.\end{aligned}$$

为什么在右端的积分前有一个$(-1)^n$呢？这是由边界的方向所确定的(也可以说是约定的)，对I^n而言，它的$(n, 0)$面的符号是$(-1)^n$。

这样便证明了在情形2时也有

$$\int_D d\omega=\int_{\partial D}\omega.$$

再由引理可知道斯托克斯公式成立。证毕。

场论中的三个基本公式都是这个定理的特殊情况。此外，设M是实直线，D是区间(a, b)，于是

$$\partial D = \{b\} - \{a\}.$$

再设ω是M上的一个0-形式，即ω是连续可微的函数f，

$$d\omega = f'(x)dx.$$

这时斯托克斯公式就是

$$\int_{(a,b)} f'(x)dx = \int_{\partial(a,b)} f = f(b) - f(a).$$

即著名的牛顿-莱布尼茨公式。由于这一原因，我们又称**斯托克斯定理是高维的微积分基本定理**。

周期

先考察一个具体情形。设

$$M = R^3 - \{0\},$$

即在3维欧几里得空间内去掉原点，M是3维微分流形，再设ω是定义在M上的一个1-形式，并且它是闭的，即$d\omega = 0$，它是不是恰当形式？由于M不是星形区域，因此不能应用5.2中的庞卡莱定理，但由斯托克斯定理知道，ω是恰当的，这是因为只要作

$$f(x, y, z) = \int_{(1,0,0)}^{(x,y,z)} \omega,$$

右端的积分沿任何不经过原点的光滑(或分段光滑)曲线C从$(1, 0, 0)$到(x, y, z)。不难证明

(i) 积分与曲线C无关，只与起点$(1, 0, 0)$和终点(x, y, z)有关；

(ii) $\omega = df$，即ω是恰当形式；

(iii) 沿单位圆周C(即以原点为中心、以1为半径的任何圆周)上的积分

$$\int_C \omega = 0.$$

并且(i)、(ii)和(iii)是等价的。

再设ω是$M = R^3 - \{0\}$上的一个2-形式，并且是闭的，情况

就比较复杂了。对M内的任何一个局部的星形区域D，由5.2的庞卡莱定理，ω在D内是恰当的。但在整体M上如何呢？如果它也是恰当的，即存在1-形式η，使$\omega=d\eta$在整个M上成立，则由斯托克斯定理，沿单位球面

$$S_2=\{(x, y, z)\in R^3|x^2+y^2+z^2=1\}$$

上的积分

$$\int_{S_2}\omega=\int_{S_2}d\eta=\int_{\partial S_2}\eta=0$$

(右端为何是0？请读者考虑)，但这并不总是成立的，因此不能说ω是恰当的。正确的结论应该是：如果ω是M上的一个2-形式，并且是闭的，又如果

$$\int_{S_2}\omega=0,$$

则ω是M上的恰当形式。这一结论作为习题请读者完成。

现在将1-形式在圆周上的积分和2-形式在2维球面上的积分这一概念加以拓广。设C是一个k-链，如果

$$\partial C=0,$$

称C是k-环，例如圆周以及和圆周同胚的闭曲线都是1-环，2维球面以及和它同胚的闭曲面都是2-环。又设ω是闭的k-形式，称ω在k-环C上的积分$\int_C\omega$是相对于k-环C的ω的周期。

由斯托克斯定理立即知道，如果k-环C是某个$(k+1)$-链b的边界，即$C=\partial b$，那么闭的k-形式(相对于C)的周期

$$\int_C\omega=\int_{\partial b}\omega=\int_b d\omega=0.$$

下面我们不加证明的给出德拉姆(De Rham)定理。

定理 设ω是一个闭形式，则ω是恰当的当且仅当它的所有周期都是0。

习　题

1. 对3-单形和4-单形S，验证

$$\partial^2 S = \partial(\partial S) = 0.$$

2. 对3维单位正方形I^3，验证

$$\partial^2 I^3 = \partial(\partial I^3) = 0.$$

3. 设S是球面，$S = \{(x_1, \cdots, x_n) \in R^n | x_1^2 + x_2^2 + \cdots + x_n^2 = 1\}$，$\omega$是一个$(n-2)$-形式，证明

$$\int_S d\omega = 0.$$

4. 证明：在一个流形M上，设ω是1-形式，如果ω的所有周期都等于0，则1-形式ω是恰当形式。

5. 设M是2维微分流形，ω是M上的1-形式，如果对M的任一坐标邻域U，ω在U内是恰当的(即存在0-形式f，使在U内有$df=\omega$成立)，称ω是局部恰当的。证明ω是局部恰当的当且仅当ω是闭的。

6. 设$M = R^3 - \{0\}$，ω是M上的2-形式，并且是闭的，又设

$$\int_{S_2} \omega = 0,$$

S_2是球面$\{(x, y, z) \in R^3 | x^2 + y^2 + z^2 = 1\}$，证明$\omega$是恰当形式。

参 考 书 目

Goffman, C.,《多元微积分》，史济怀等译，人民教育出版社，(1979年).

斯皮瓦克，《流形上的微积分》齐民友、路见可译，科学出版社，(1981年).

Schreiber, M 《微分形式导论》，白正国译，人民教育出版社，(1981年).

弗列明，《多元函数》，庄亚栋译，人民教育出版社，(1982年).

Flander, H., *Differential Forms with Applications to Physical Sciences*, Academic Press, New York, London(1963).

Woll, W., *Functions of Several Variables*, Harcourt, Brace and World Inc, (1966).

欧阳光中、朱学炎、秦曾复，《数学分析》，上海科学技术出版社，(1982年).

索　引

一画～三画

四　画

五　画

六　画

七　画

八 画

九 画

十 画

十 一 画

十 二 画

十 三 画

十 四 画 以 上